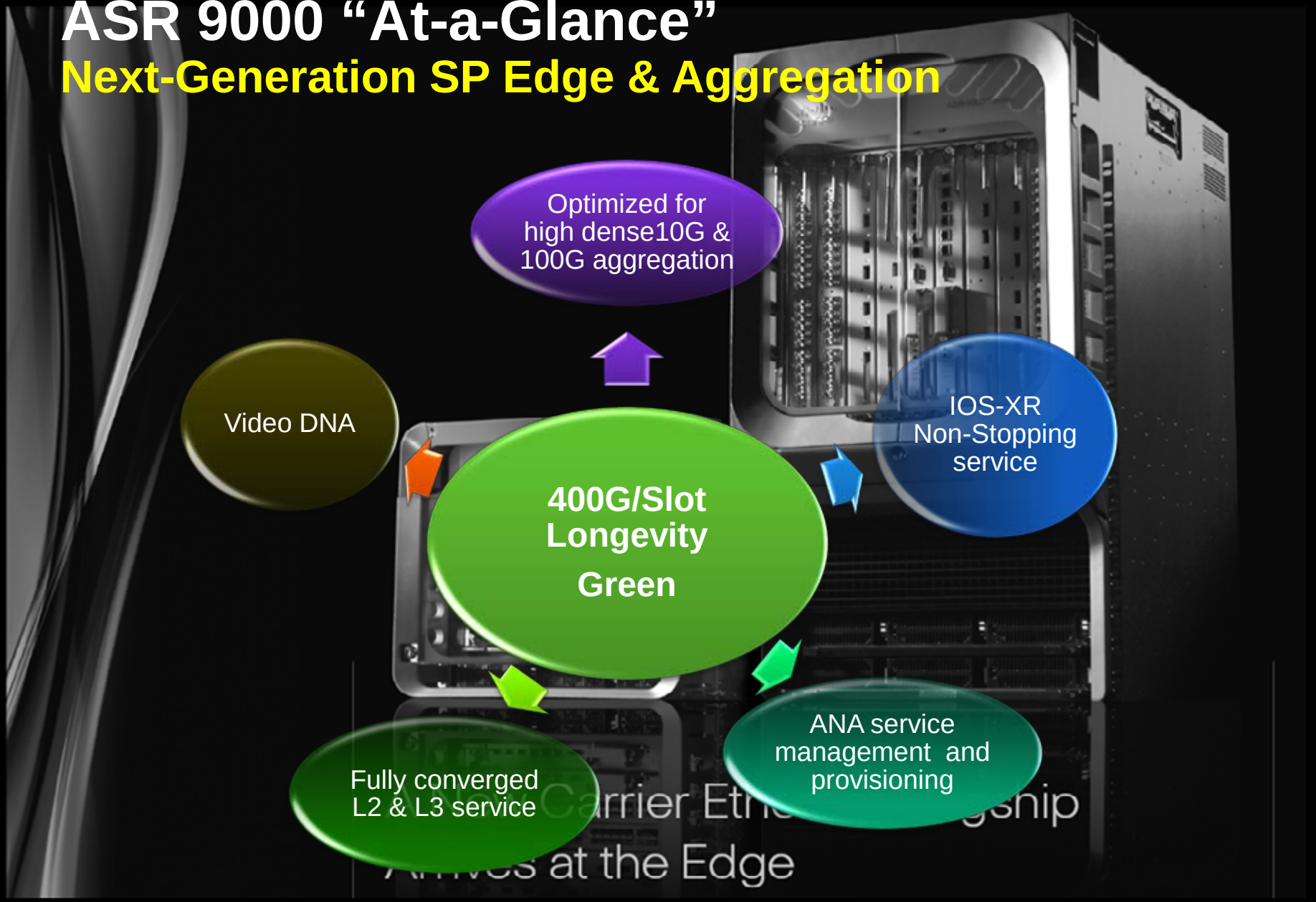
The Cisco ASR 9000 Architecture

BRKARC-2003

ASR 9000 “At-a-Glance”

Next-Generation SP Edge & Aggregation



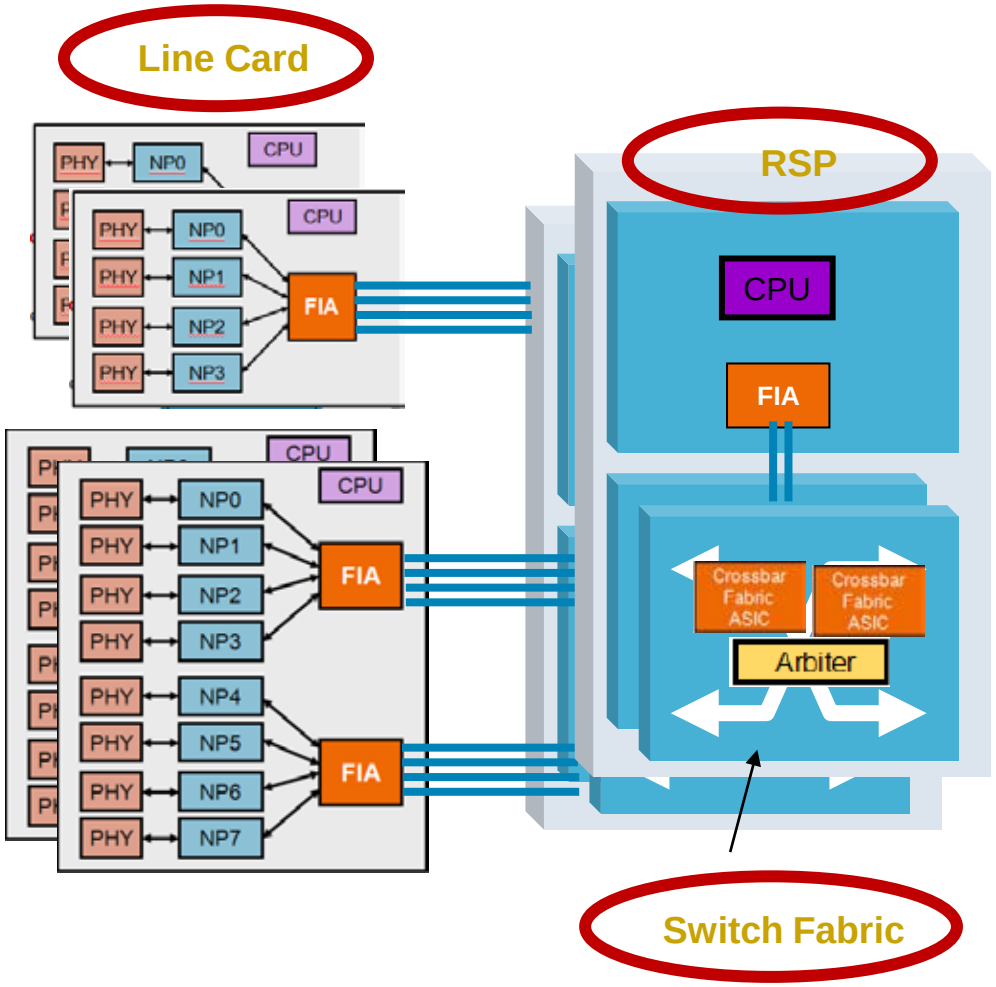
ASR 9000 System Architecture “At-a-Glance”

Fully Distributed Architecture for High Performance and High Multi-dimensional Control Plane Scale

- Data forwarding is fully distributed on the line cards
- Control plane split among RSP and LC CPU (same type of CPU as RSP)
- L2 protocols, BFD, CFM, Netflow runs on the LC CPU for high scale

True Modular OS for HA and Operational Simplicity

- Micro-kernel based, true modular OS
- High availability and System stability
- SW patch granularity for operational simplicity

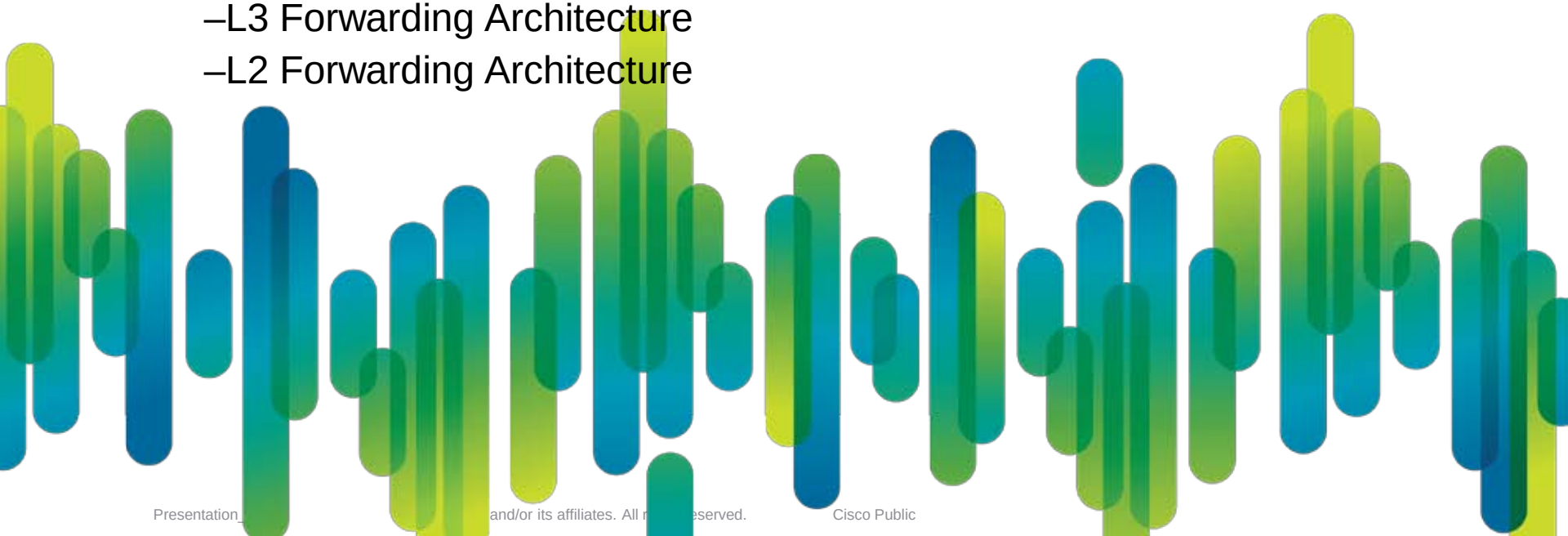


Active-Active Switch Fabric
Guarantee “0” packet loss during RSP failover

© 2010 Cisco and/or its affiliates. All rights reserved.

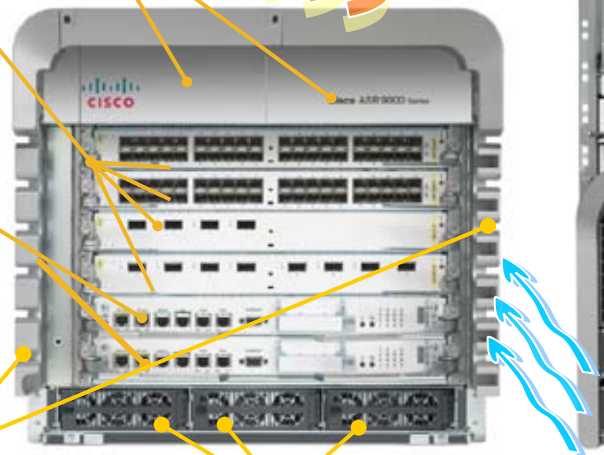
Agenda

- **Hardware Architecture**
  - –Chassis
  - –RSP
  - –Line Cards
  - –Switch Fabric
  - –System QoS
  - –Multicast
- **Software Architecture**
  - –IOS-XR and High Availability
  - –L3 Forwarding Architecture
  - –L2 Forwarding Architecture



ASR 9010 and ASR 9006 Chassis

Identical HW components across two chassis*



Front-to-back airflow

Integrated cable management with cover

System fan trays

Side-to-back airflow

Line Card (0-3)

RSP (0-1)

cable management

Three Modular Power Supplies

RSP (0-1)

Line Card (0-3, 4-7)

System fan trays

Air draw

Six Modular Power Supplies

* Fan tray is different

© 2010 Cisco and/or its affiliates. All rights reserved.

Power and Cooling

Existing Power Supply and Fan are ready for 400G/slot

ASR 9010 Fan Tray

ASR 9006 Fan Tray

- Fans unique to chassis
- Variable speed for ambient temperature variation
- Redundant fan-tray
- Low noise, OSHA compliant

Power Supply

DC Supplies

A
B
2.1/1.5 kW

A
B
2.1 kW

AC Supplies

A
3 kW

B
3 kW

- 6 & 10 slot use same power supplies
- Single power zone
- All power supplies run in active mode
- Power draw shared evenly
- 50 Amp DC Input or 16 Amp AC for Easy CO Install

© 2010 Cisco and/or its affiliates. All rights reserved.

Power Check and Rules

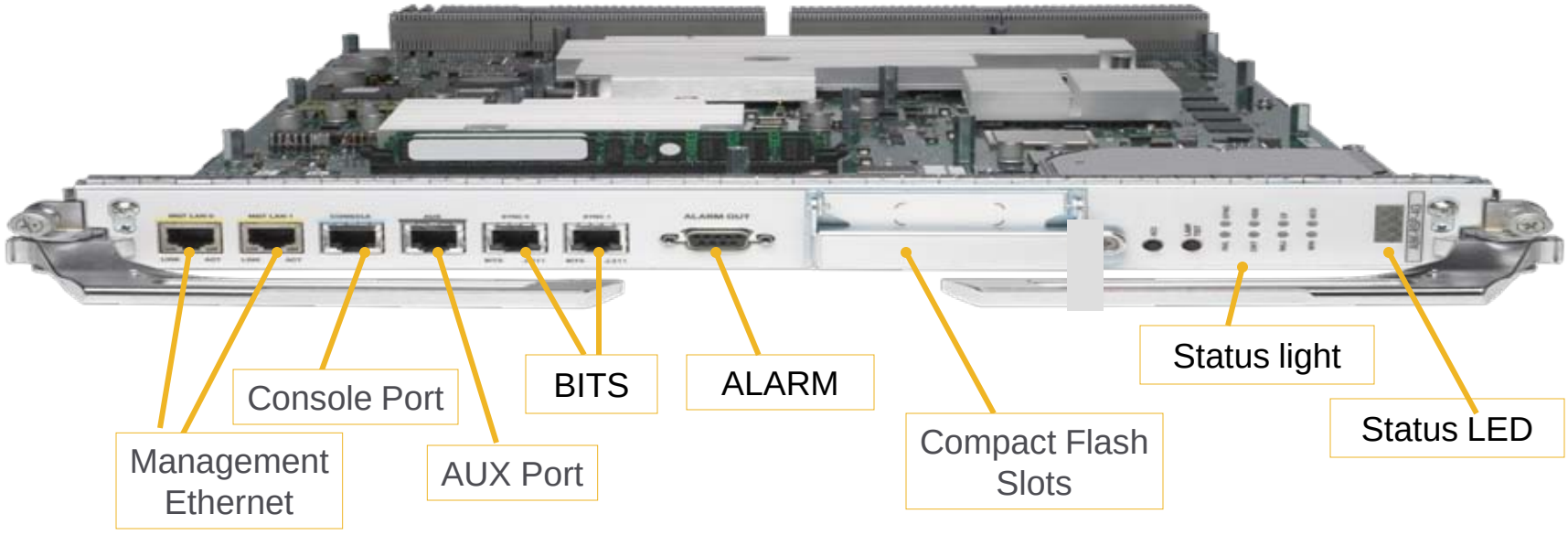
- Available power is checked when:
  - –An LC card is inserted
  - –An LC card is powered up via the CLI
  - –An LC card is reset via “hw-mod reload”
- If the system does not have enough available power to accommodate the LC, then the LC becomes “UNPOWERED”
- Installing new power supplies will *not* automatically power up any UNPOWERED line cards. The user can force a recheck using:

  “hw-mod reload loc <>”
- RSP and Fan Tray cards are given priority allocation of power budget
- LC power budget is checked in numeric order until it is exhausted. The actual power up of the LC's is in parallel, and asynchronous

© 2010 Cisco and/or its affiliates. All rights reserved.

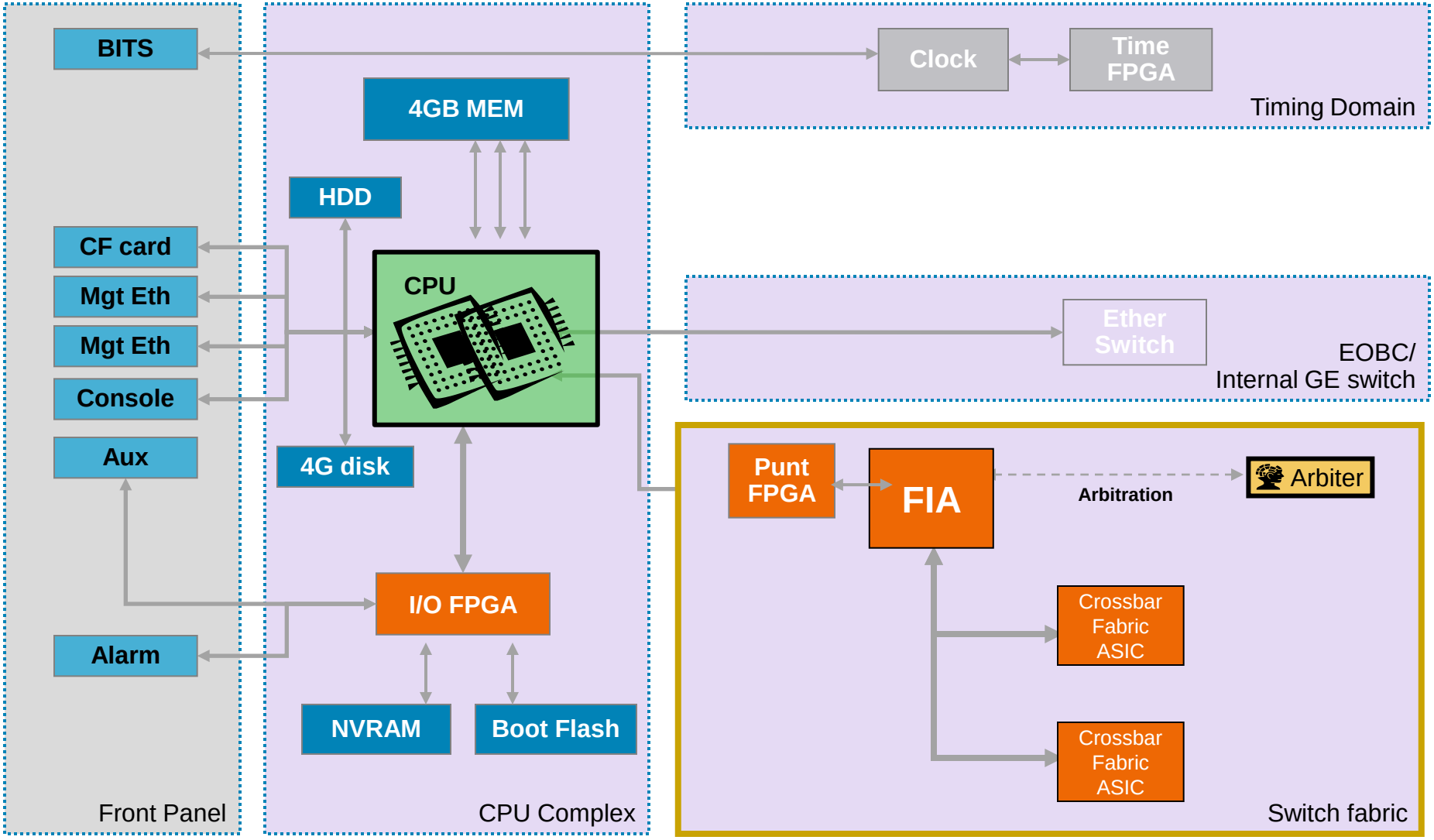
RSP Engine

- Performs control plane and management functions
- Dual Core CPU processor with 4GB DRAM
- 2MB NVRAM, 4GB internal bootdisk, 2 external compact flash slots
- Dual Out-of-band 10/100/1000 management interface
- Console & auxiliary serial ports
- Hard Drive: 70G HDD



© 2010 Cisco and/or its affiliates. All rights reserved.

RSP Engine Architecture



© 2010 Cisco and/or its affiliates. All rights reserved.
Cisco Public

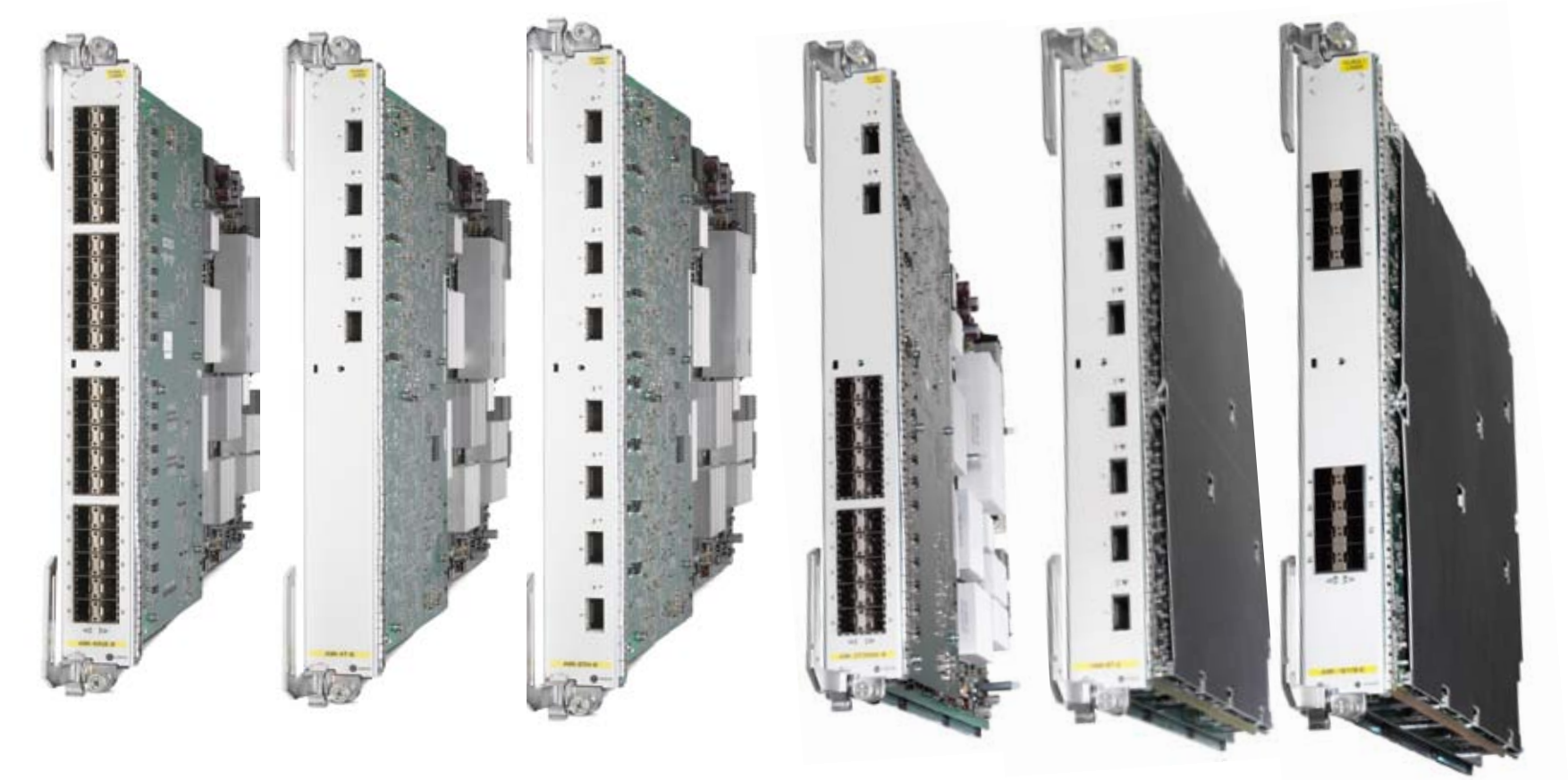
Agenda

- **Hardware Architecture**
  - –Chassis
  - –RSP
  - –Line Cards
  - –Switch Fabric
  - –System QoS
  - –Multicast
- **Software Architecture**
  - –IOS-XR and High Availability
  - –L3 Forwarding Architecture
  - –L2 Forwarding Architecture

© 2010 Cisco and/or its affiliates. All rights reserved.

Ethernet Line Card Family

- Common HW ASIC and HW architecture
- Identical SW features
- Line card Memory options for QoS scale



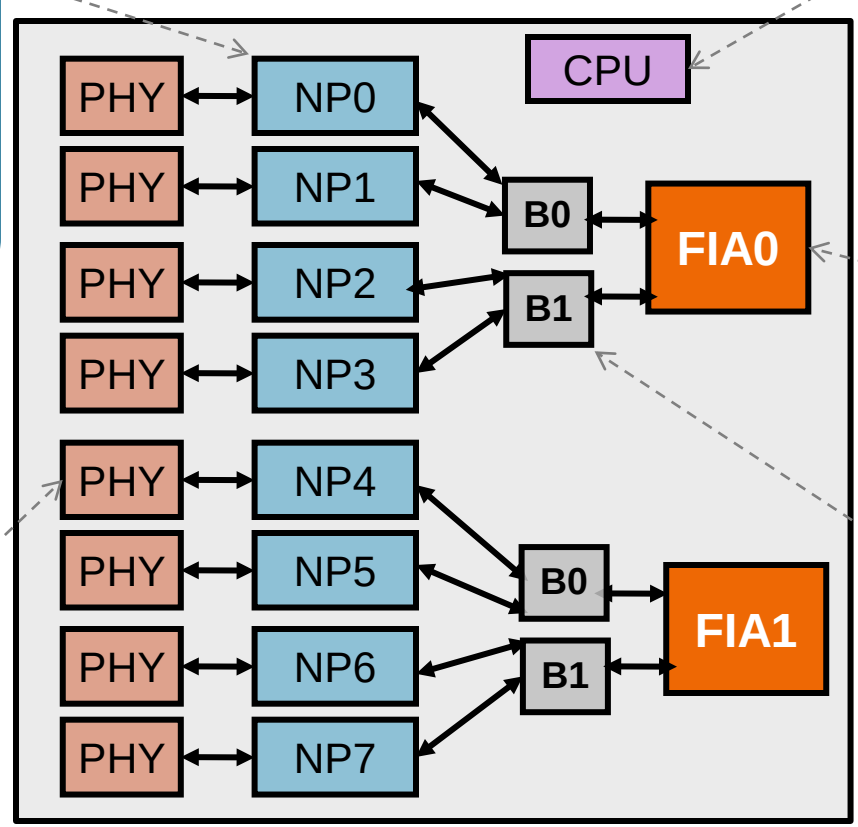
A9K-40G | A9K-4T | A9K-8T/4 | A9K-2T20G | A9K-8T | A9K-16T/8

© 2010 Cisco and/or its affiliates. All rights reserved.

Line Card Architecture – Hardware Components

NP: Network Processor
Main forwarding ASIC
L2 & L3 forwarding, features (QoS, ACL, etc), control plane policing, mcast replication, etc

10Gbps bi-directional with features applied
30Gbps raw bandwidth



CPU (same as RSP)
Program HW forwarding tables
Distributed Control planes
SW switched packets
Inline Netflow

FIA: Fabric Interface ASIC
Provide non-blocking data connection to switch fabric
Internal system queues/VoQ
Intelligent mcast replication

10G PHY for one 10G port, or 10x1G port

B: Bridge FPGA
Provide non-blocking data connection between NP and FIA
Internal System queues
Intelligent mcast replication

Example: A9K-8T

Note, Bridge FPGA provide non-blocking connection between NP and the FIA. Functionally it does the HW conversion due to different interface format on NP and FIA. It's part of the switch fabric connection. To make it logically simple, it will be removed from the remaining slides.

© 2010 Cisco and/or its affiliates. All rights reserved.

Line Card HW Components – Counters

- RP/0/RSP1/CPU0:SJC#show controllers fabric ?

  Arbiter Arbitration ASIC show screens.

  Crossbar XBAR ASIC show screens.

  fia Show command for fabric interface asic

- RP/0/RSP1/CPU0:SJC#show controllers fabric **fia bridge** stats location 0/0/cpu0 **B0**

- RP/0/RSP1/CPU0:SJC#show controllers fabric **fia** stats location 0/0/cpu0 **FIA**

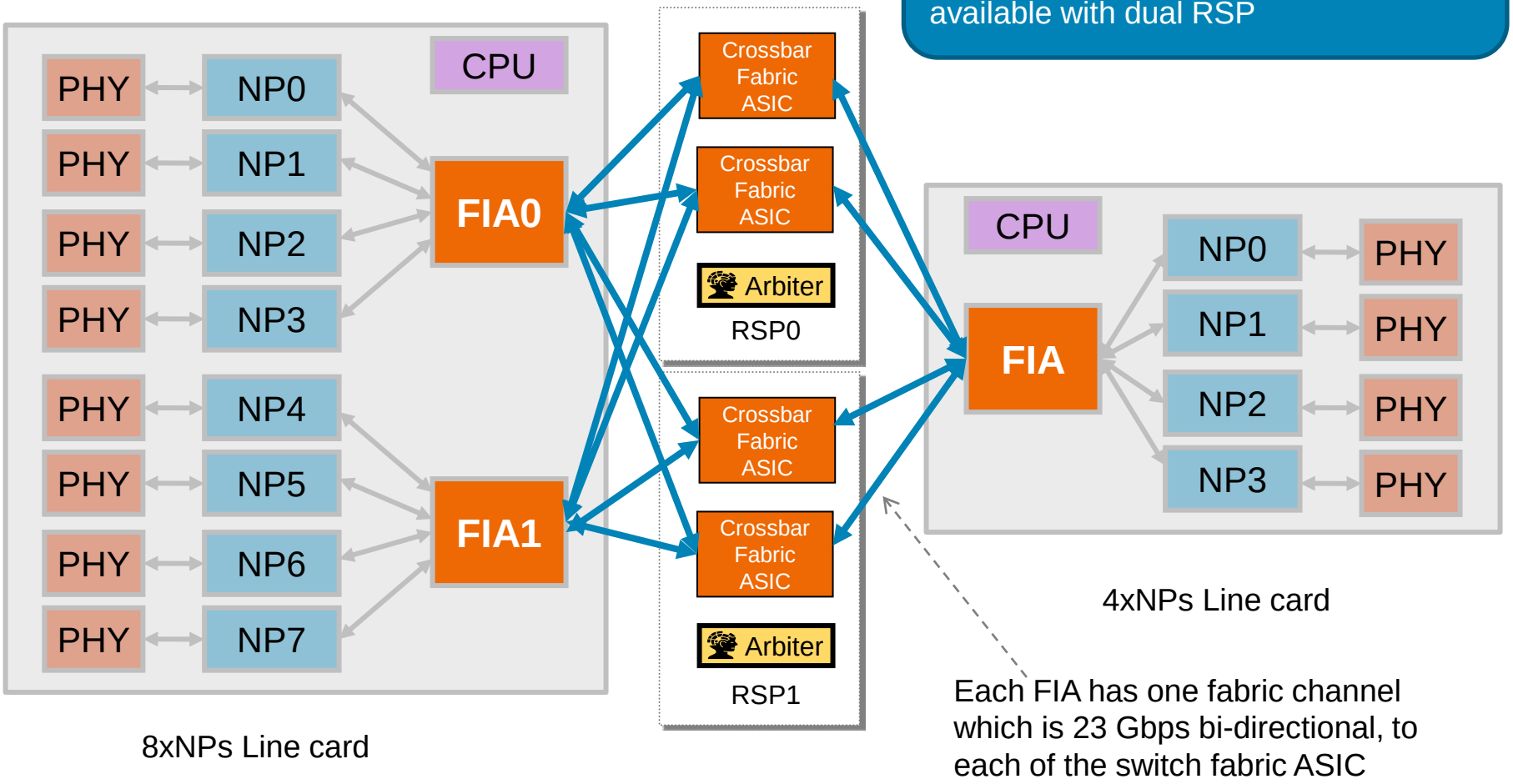
- RP/0/RSP1/CPU0:SJC#show controllers **np** ? NP

```
counters          Display contents of global stats counters
 crashinfo        Display NP Crash info
 drvlog           Display Driver Logging
 fabric-counters  XAUI counters dump
 interrupts       Show NP interrupt data
 memory           NP Raw Memory Dump
 portMap          Show port mapping on NP
 ports            Shows physical ports associated with each np
<snip>
```


© 2010 Cisco and/or its affiliates. All rights reserved.

Line Card Architecture – Fabric Connection

- Non-blocking line-rate fabric bandwidth to each LC with single RSP
- Extra (double) fabric bandwidth available with dual RSP



8xNPs Line card

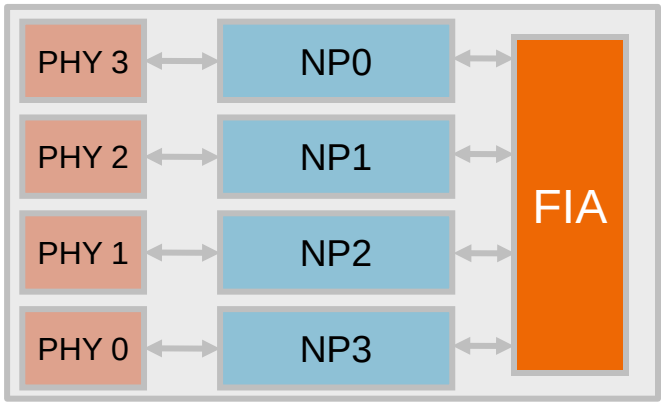
4xNPs Line card

Each FIA has one fabric channel which is 23 Gbps bi-directional, to each of the switch fabric ASIC

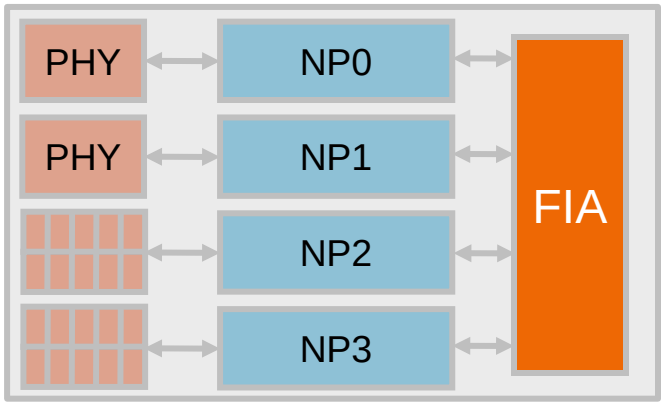
© 2010 Cisco and/or its affiliates. All rights reserved.

4xNPs Line Card Family

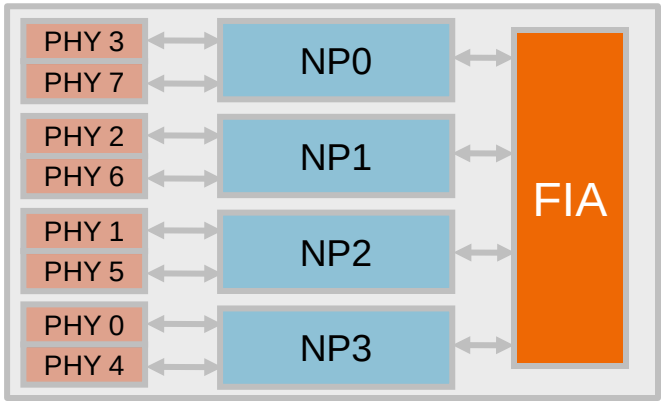
A9K-4T-E/B/L



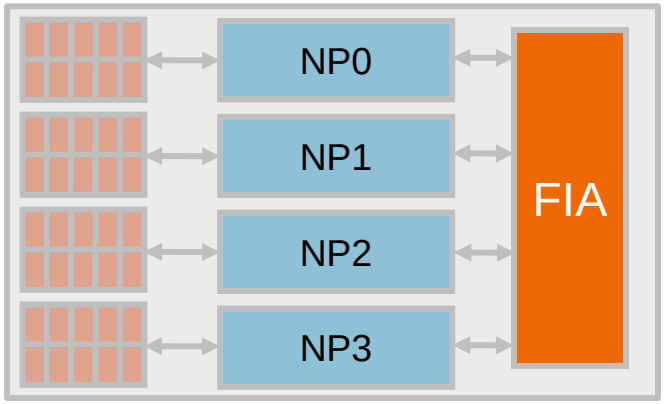
A9K-2T20G-E/B/L



A9K-8T/4-E/B/L



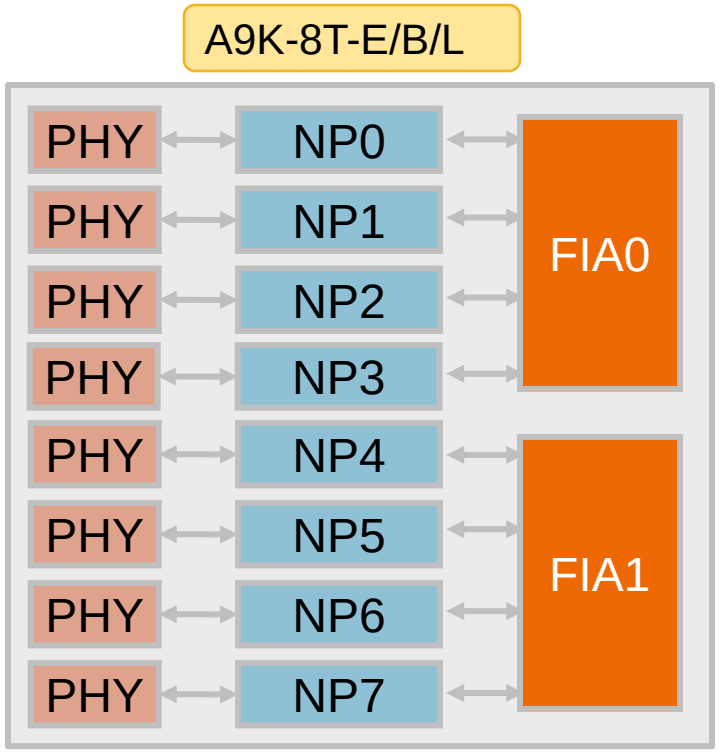
A9K-40G-E/B/L

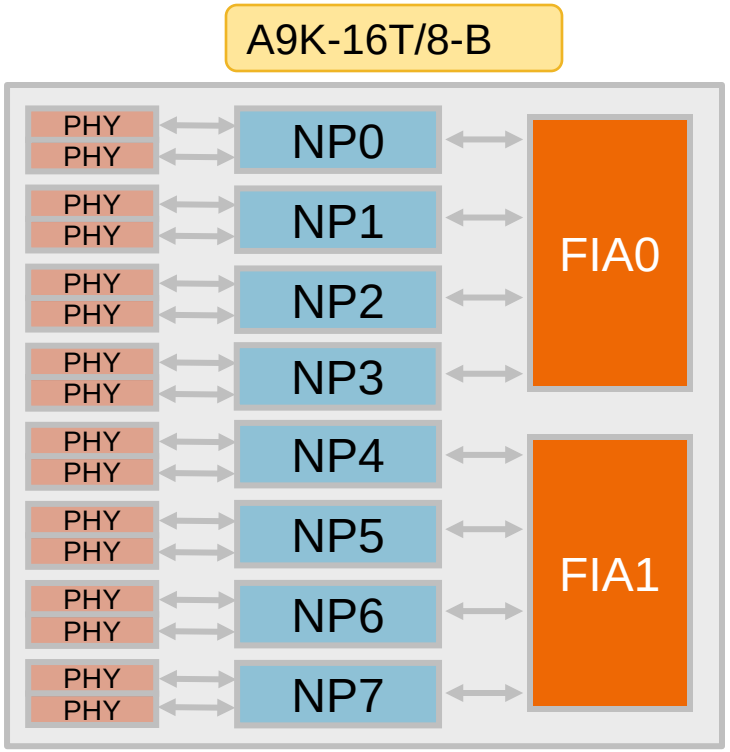


Oversubscribed line card
Up to 60Gbps bandwidth

© 2010 Cisco and/or its affiliates. All rights reserved.
Cisco Public

8xNPs Line Card Family

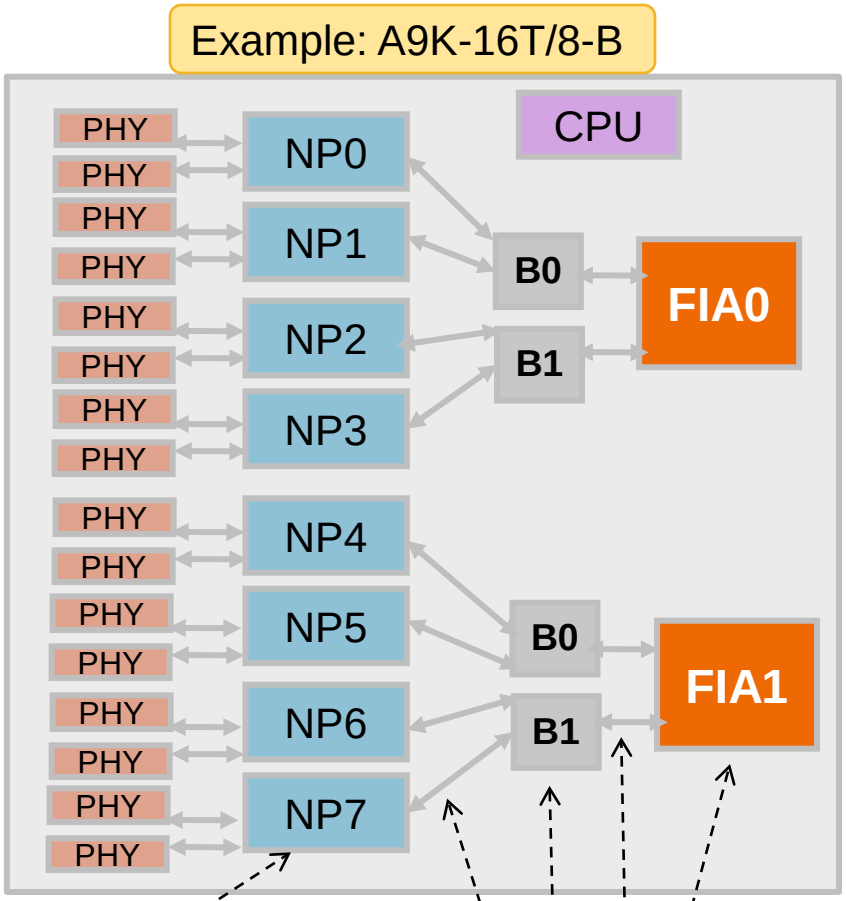




Oversubscribed line card
Up to 120Gbps bandwidth

© 2010 Cisco and/or its affiliates. All rights reserved.
Cisco Public

Line Card Architecture – Internal Bandwidth
Extra Bandwidth for Oversubscribed LCs

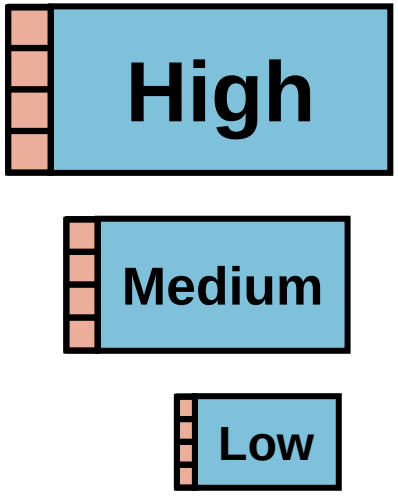


A9K-16T/8-B
120Gbps raw bandwidth

A9K-8T/4
60Gbps raw bandwidth

© 2010 Cisco and/or its affiliates. All rights reserved.
Cisco Public

Line Card Memory Options – Queue Scale

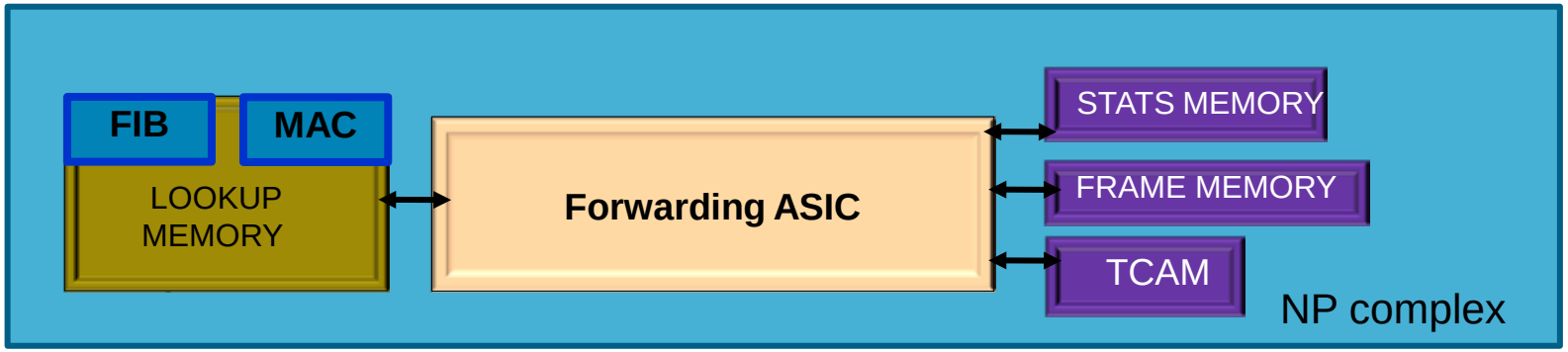


- 3 memory options for each line card: **E**xtended (or high queue), **B**ase (medium queue), **L**ow (low queue)*
- Different memory option has different QoS queue scale and L2 sub-interface scale. All other system wide scale is the same across different type of the line cards, including FIB, MAC address, Bridge-domain, L3 sub-interface, VRF, etc
- All line cards have the same HW → Identical features
- Mixed different type of line cards are supported on the same chassis with same system wide scale and identical features

* A9K-16T/8 only have “B” option. It doesn’t have “E” or “L” option as of the 3.9.1 release

© 2010 Cisco and/or its affiliates. All rights reserved.

L/B/E Line Cards – What's the Difference?



- Each NPU has Four Main Associated memories TCAM , Search/Lookup memory , Frame/buffer memory and statistics memory
  - –TCAM is used for VLAN tag, QoS and ACL classification
  - –Lookup Memory is used for storing FIB tables, Mac address table and Adjacencies
  - –Stats memory is used for all interface statistics, forwarding statistics etc
  - –Frame memory is buffer memory for Queues
- E/B/L line card have different TCAM , Stats and Frame Memory size, which give different scale number of the QoS queues and L2 sub-interfaces per line card
- Lookup Memory is the same across line card s→ why?
  - –To support mix of the line cards without impacting the system wide scale including routing, multicast, MAC address, L3 interface, MPLS label space scale

© 2010 Cisco and/or its affiliates. All rights reserved.

Optics

Standard, CWDM, & DWDM XFPs/SFPs/SFP+ available
IPoDWDM G.709 FEC/EFEC

GE - SFP Optics
T, S, L, & Z

TenGE – XFP Optics
LR, ZR, & ER

© 2010 Cisco and/or its affiliates. All rights reserved.

SIP/SPA – Non-Ethernet Support

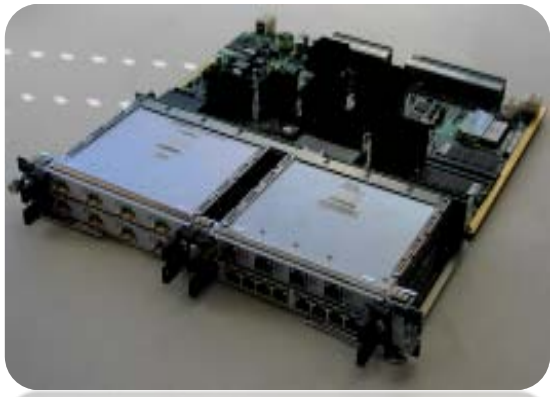
Strong QoS
- 128+K Queues,
- 128+K Policers
- H-QoS
- Color Policing

Scalability
- Distributed control and data plane; 20G; 4 Bays
- L3 i/f, route, session, protocol – scaled up for MSE needs

High Availability
- IC-Stateful Switch Over capability,
- MR-APS
- IOS-XR base for high scale and reliability

Powerful and Flexible QFP Processor
- Flexible ucode Architecture for Feature Richness
- L2 + L3 Services: FR, PPP, HDLC, MLPPP, LFI, L3VPN, MPLS, Netflow, 6PE/VPE

A9K-SIP-700



© 2010 Cisco and/or its affiliates. All rights reserved.

*Not Commtted

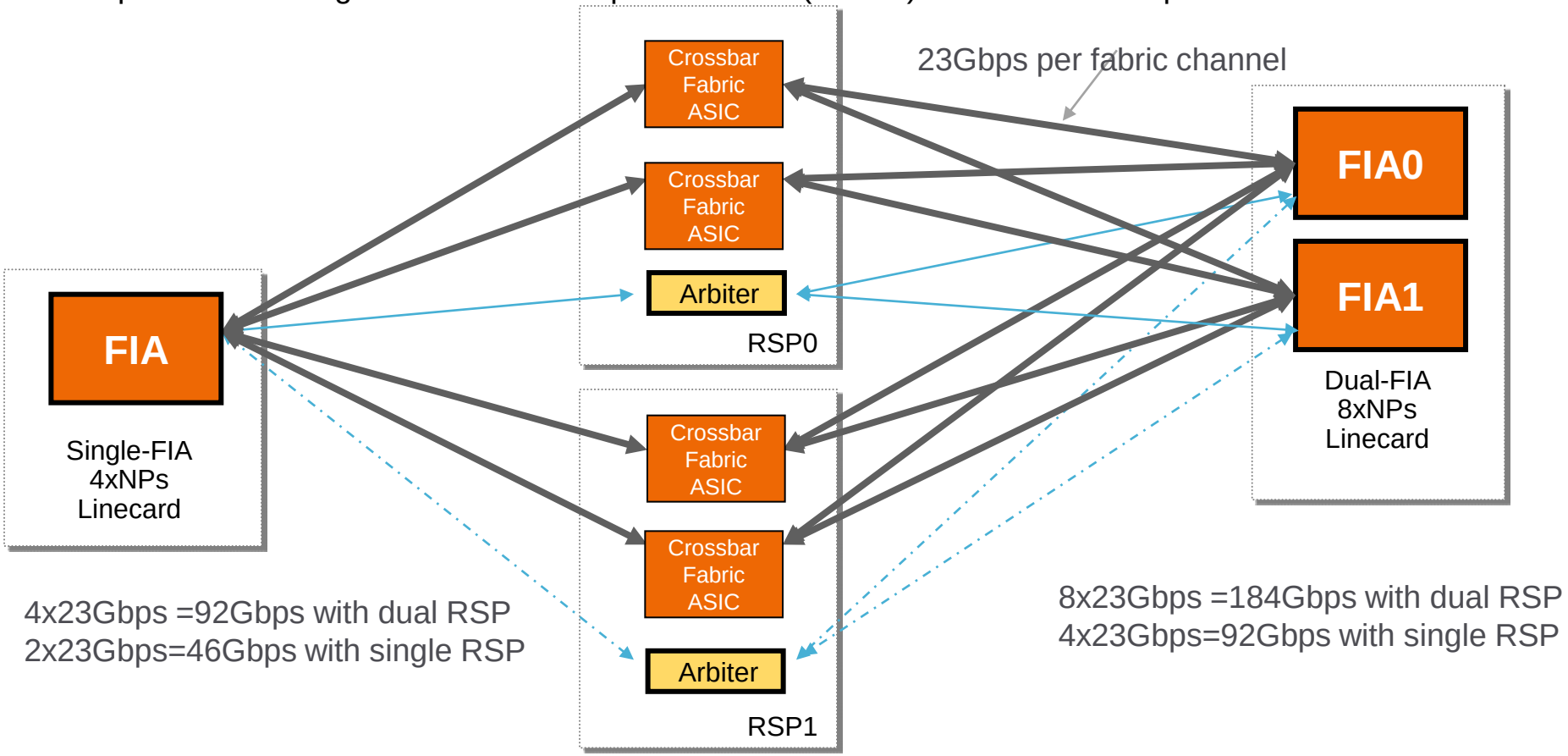
Agenda

- **Hardware Architecture**
  - –Chassis
  - –RSP
  - –Line Cards
  - –Switch Fabric
  - –System QoS
  - –Multicast
- **Software Architecture**
  - –IOS-XR and High Availability
  - –L3 Forwarding Architecture
  - –L2 Forwarding Architecture

© 2010 Cisco and/or its affiliates. All rights reserved.

Switch Fabric Overview

- Switch fabric physically resides on RSP, but operate separately from RSP
- Separate data and arbitration paths
- All fabric data channels run in active mode. Arbiter is in active/standby mode, which is controlled by low level hardware signalling
- Each LC has one or two fabric interface ASIC (FIA), which provide line rate bandwidth per slot with single RSP. Dual RSP provide extra (double) fabric bandwidth per slot

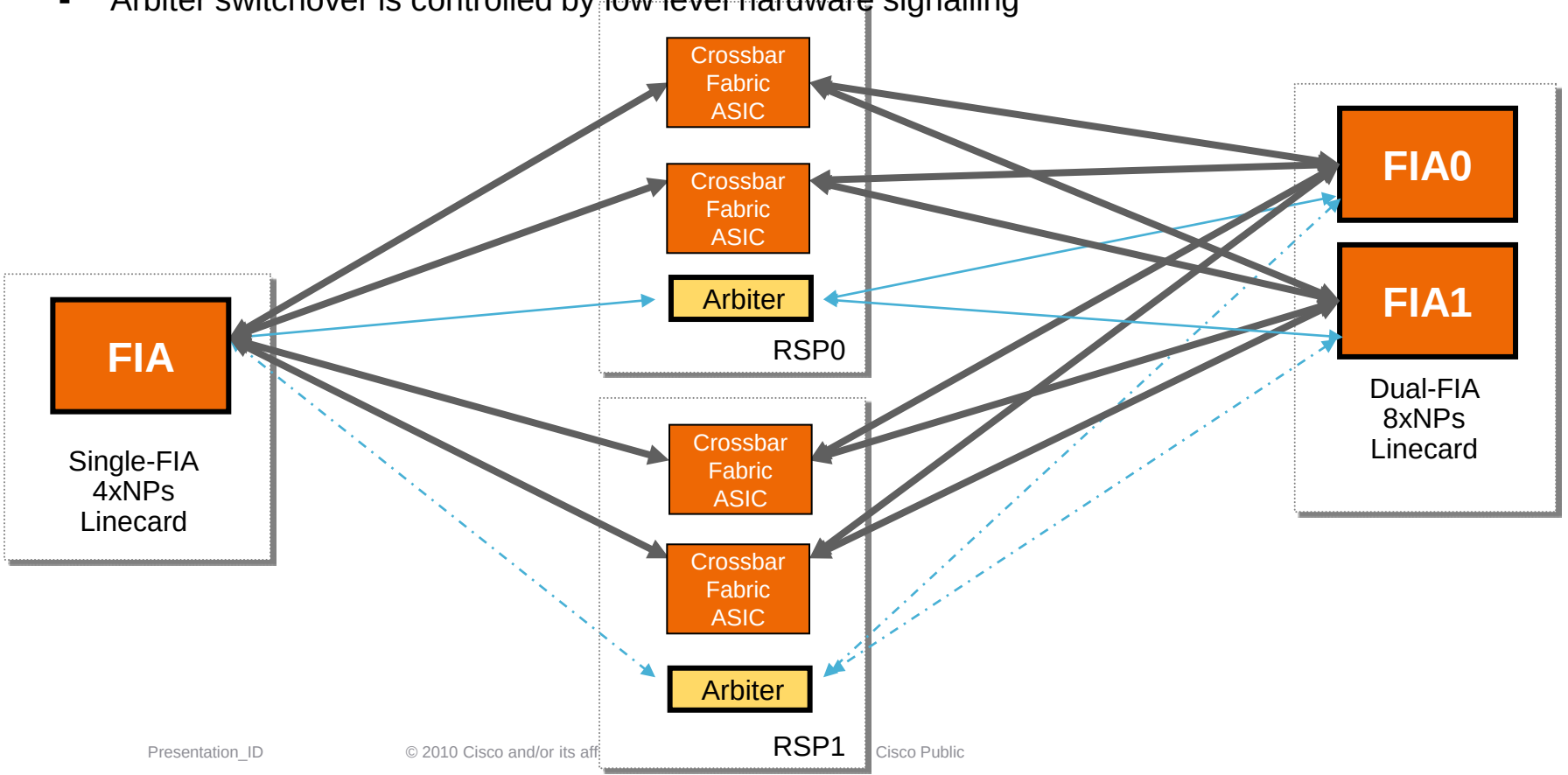


© 2010 Cisco and/or its affiliates. All rights reserved.
Cisco Public

Fabric Redundancy

Guarantee “0” packet loss during RSP failover and OIR

- Separate data and arbitration(control) paths
- All fabric data channels run in active mode for extra fabric bandwidth and instant fabric switch over
- Both Arbiters work in parallel – both answer all requests, FIAs follow active Arbiter, FIAs switch to backup if needed → instant control path switchover
- Arbiter switchover is controlled by low level hardware signalling

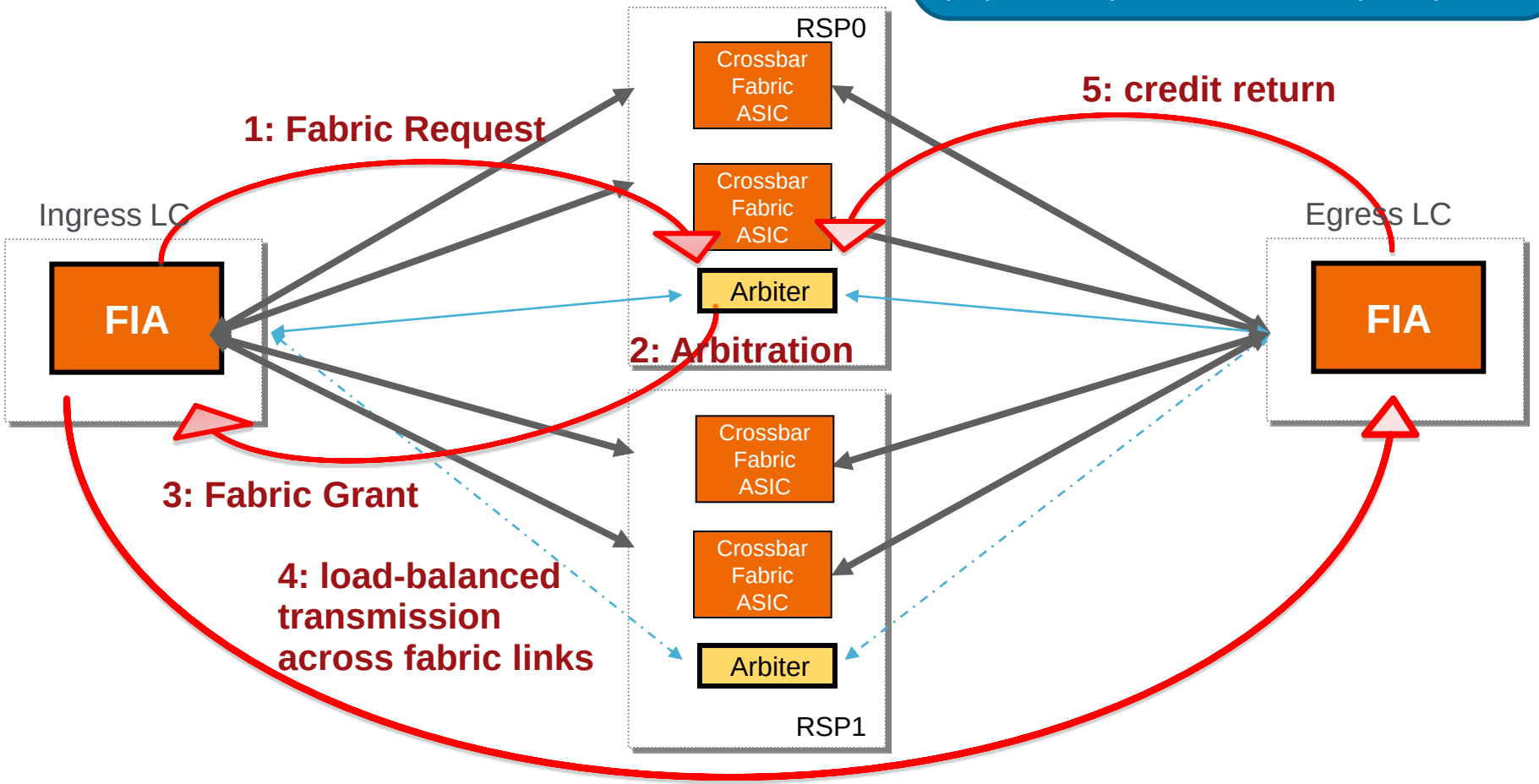


© 2010 Cisco and/or its aff

Switch Fabric Bandwidth Access Overview

Intelligent Fabric and Internal System QoS

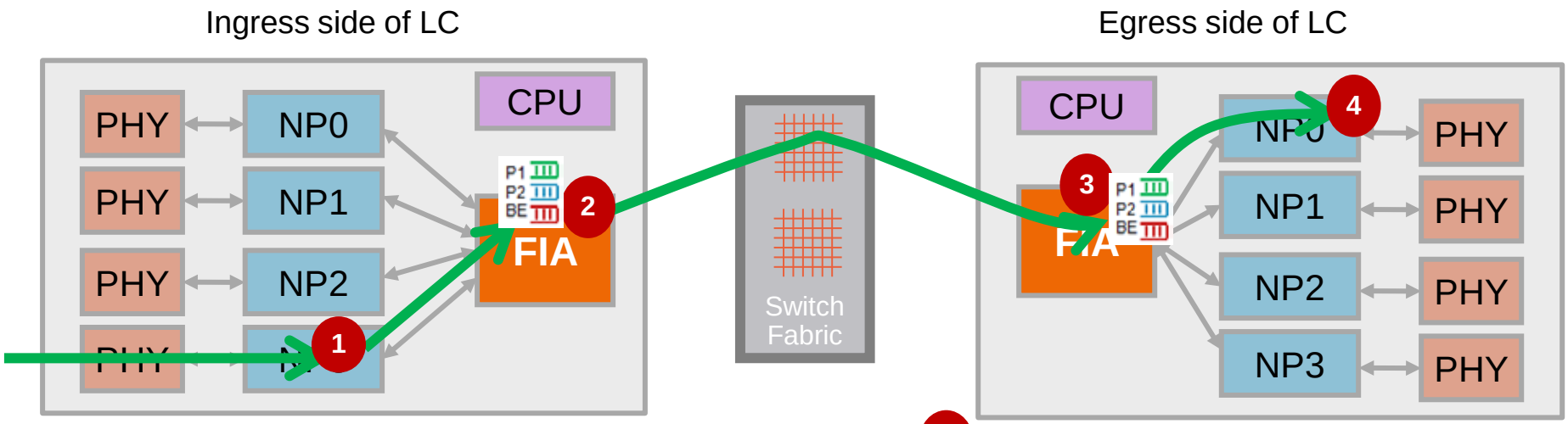
- 3 strict priority scheduling/queueing
- Back pressure and virtual output queue
- Multicast and Unicast separation (separated queues and fabric plane)



© 2010 Cisco and/or its affiliates. All rights reserved.
Cisco Public

End-to-End System QoS Queuing

End-to-End priority (P1,P2, Best-effort) propagation → Guarantee bandwidth, low latency for high priority traffic at any congestion point
3 strict priority level across all internal HW components



1 Ingress (sub-)interface QoS Queues

Configure with
Ingress MQC
4-layer hierarchy
Two strict high priority + Normal priority

2 Virtual Output Queues

Implicit Configuration
Two strict high priority + Normal priority

3 Egress FIA Queues

4 Egress (sub-)interface QoS Queues

Configure with
Egress MQC
4-layer hierarchy
Two strict high priority + Normal priority

© 2010 Cisco and/or its affiliates. All rights reserved.
Cisco Public

Backpressure and VoQ Mechanism

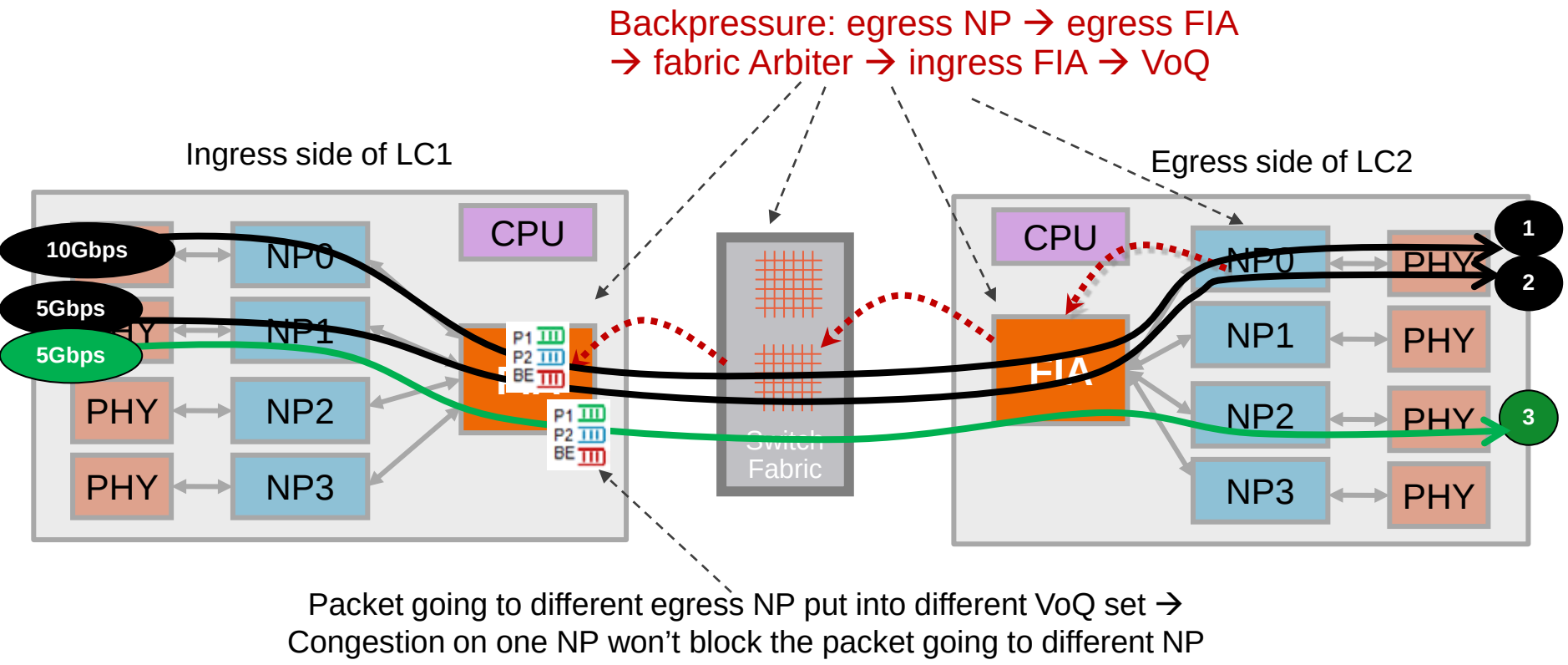
VoQ Scale: Each FIA has P1/P2/BE queue set for every NP and RSPs in the entire system

Egress NP congestion → → backpressure to ingress FIA →

Packet is en-queued in the dedicated VoQ →

No impact of the packet going to different egress NP →

No head-of-line-block issue



© 2010 Cisco and/or its affiliates. All rights reserved.

CLI – System Queues Counters

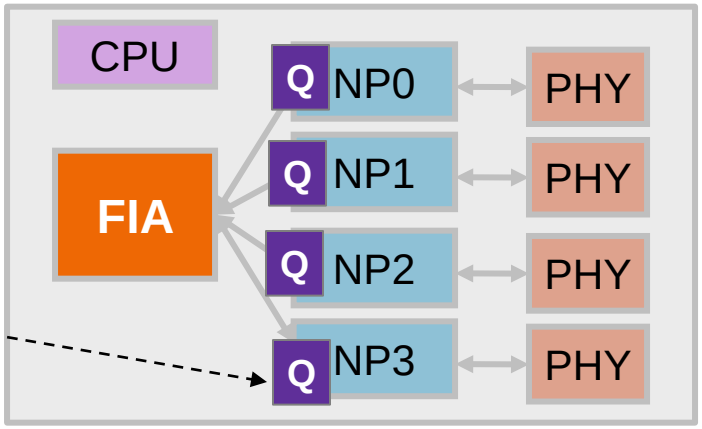
```
RP/0/RSP0/CPU0:PE2-asr#show controllers fabric crossbar ? ← switch fabric
  readreg     Read register
  serdes      Display chico info
  statistics  Packet Statistics ← switch fabric drop counters


RP/0/RSP0/CPU0:PE2-asr#show controllers fabric fia ? ← fabric interface ASIC
  bridge      show command for bridge asic
  drops       Drops for fabric interface asic ← drop counters
  errors      Errors for fabric interface asic
  link-status  Link status between fabric/xbar and fabric interface asic
  q-depth      Ingress Queue Depth Counters
  stats       Packet statistics for fabric interface asic ← packet statistics
  trace       Show controllers fia trace data

RP/0/RSP0/CPU0:PE2-asr#show controllers fabric fia bridge ? ← bridge FPGA
  ddr-status    DDR staus between bridge and fia
  flow-control  Flow control information for bridge ← flow control counters
  stats         Packet statistics for bridge ← packet statistics
  sync-status   Bridge and NP sync status
```


© 2010 Cisco and/or its affiliates. All rights reserved.

ASR 9000 (sub-)interface Level QoS Overview

User configure (sub-)interface level QoS via MQC CLIs

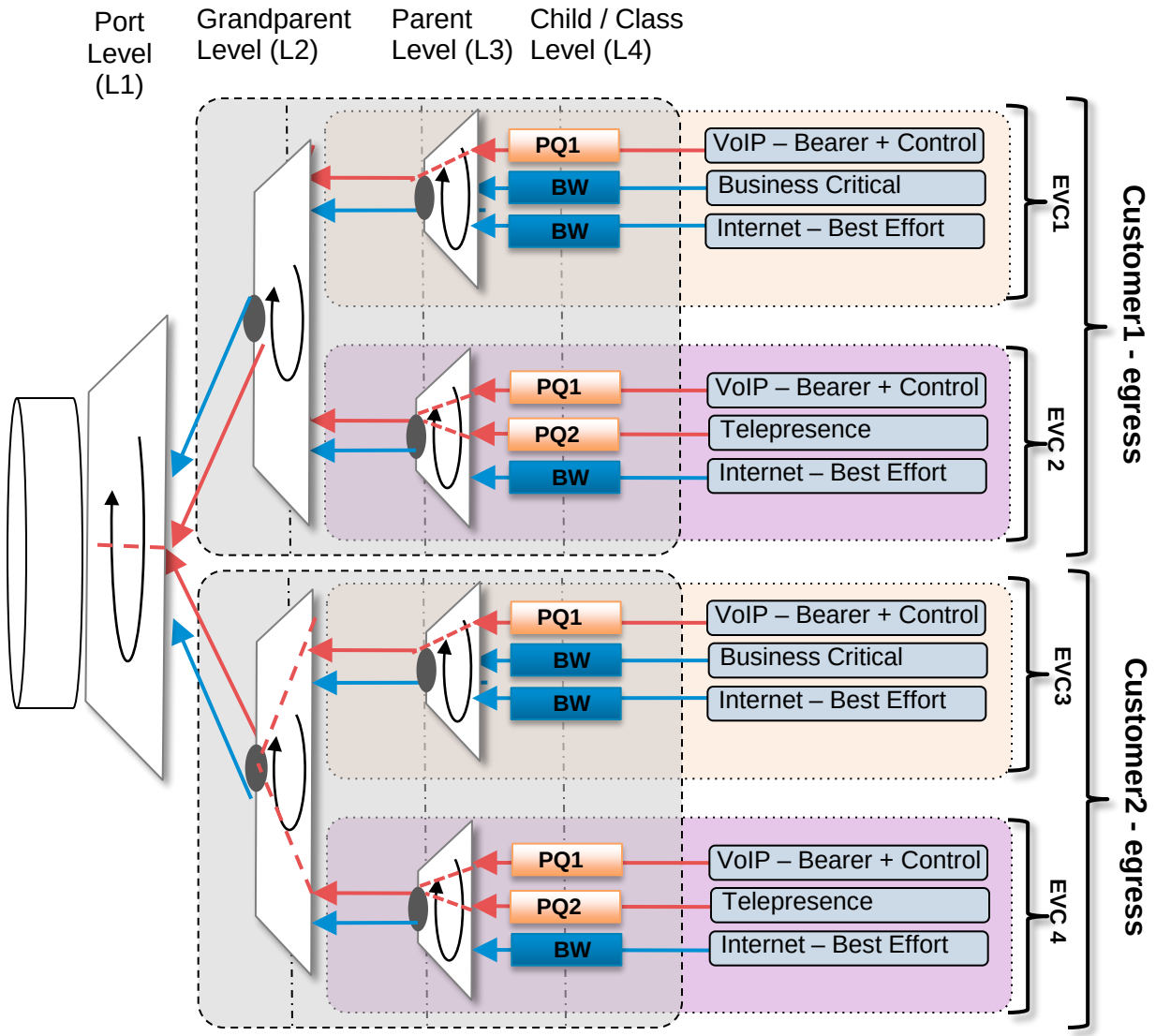
QoS function is done by special Queuing HW chip within the NP complex



- High scale
  - –Up to 3 Million queues per system (with E linecard)
  - –Up to 2 Million policers per system (with E linecard
- Highly flexible: 4 layer hierarchy queuing/scheduling support
  - –Four layer scheduling hierarchy →Port, Subscriber Group, Subscriber, Class
  - –Egress & Ingress, shaping and policing
- Three strict priority scheduling with priority propagation
- Flexible & granular classification, and marking
  - –Full Layer 2, Full Layer 3/4 IPv4, IPv6

© 2010 Cisco and/or its affiliates. All rights reserved.

4 Layer Hierarchy QoS Overview
Ingress & Egress



Note: We count hierarchies as follows:
4L hierarchy = 3 Level nested p-map
3L hierarchy = 2 level nested p-map

L1 level is not configurable but is implicitly assumed

Hierarchy levels used are determined by how many nested levels a policy-map is configured for and applied to a given (sub-) interface

© 2010 Cisco and/or its affiliates. All rights reserved.

Intelligent Multicast Replication

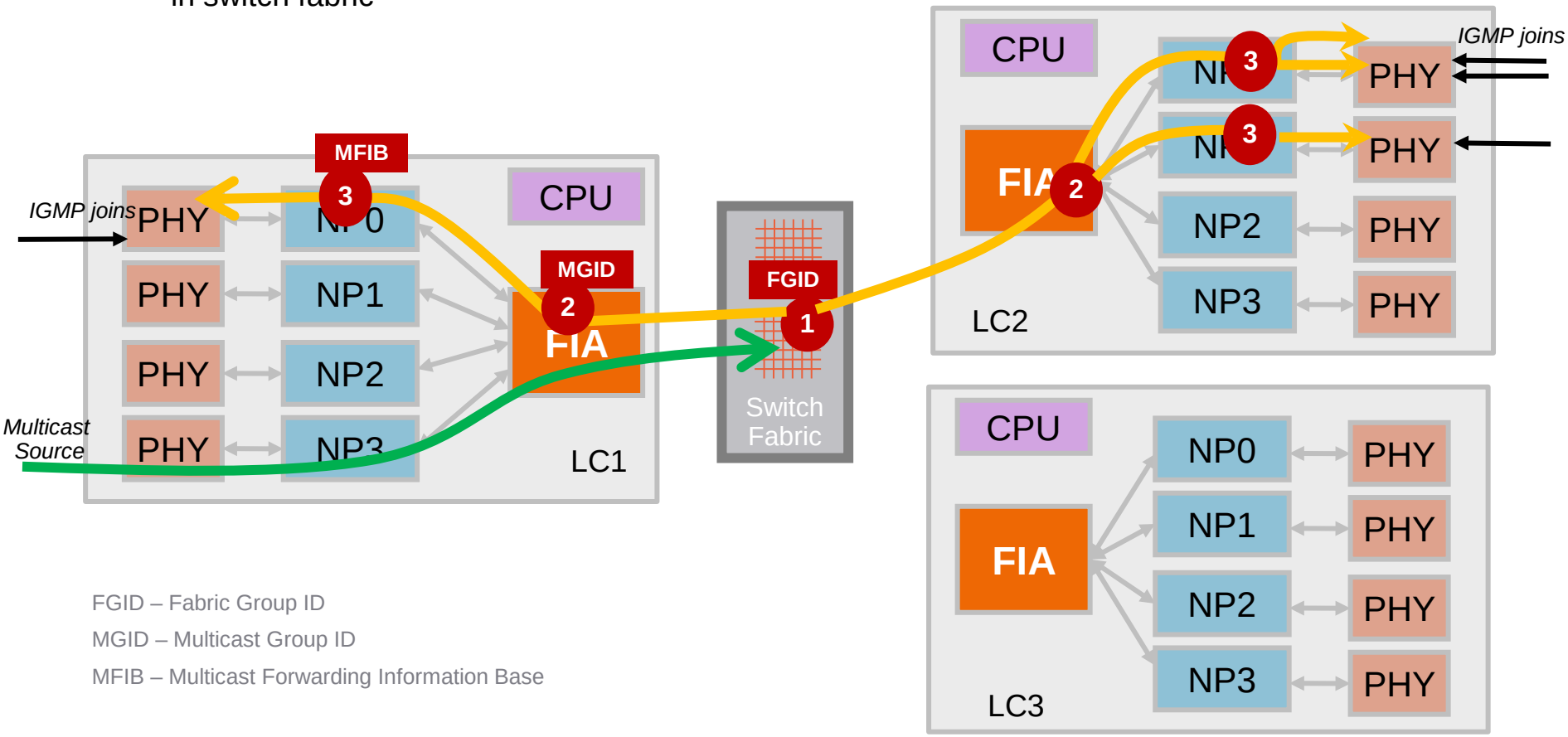
Switch Fabric and Egress LC Replication

Efficiency
Line rate
Scalable
Simple

1 **Fabric Replication** → replicate single copy to LCs which receive IGMP join, based on FGID table in switch fabric

2 **FIA Replication** → replicate single copy to NPs which receive IGMP join, based on MGID table in FIA

3 **NP Replication** → replicate copy per receiver based on multicast FIB table



FGID – Fabric Group ID

MGID – Multicast Group ID

MFIB – Multicast Forwarding Information Base

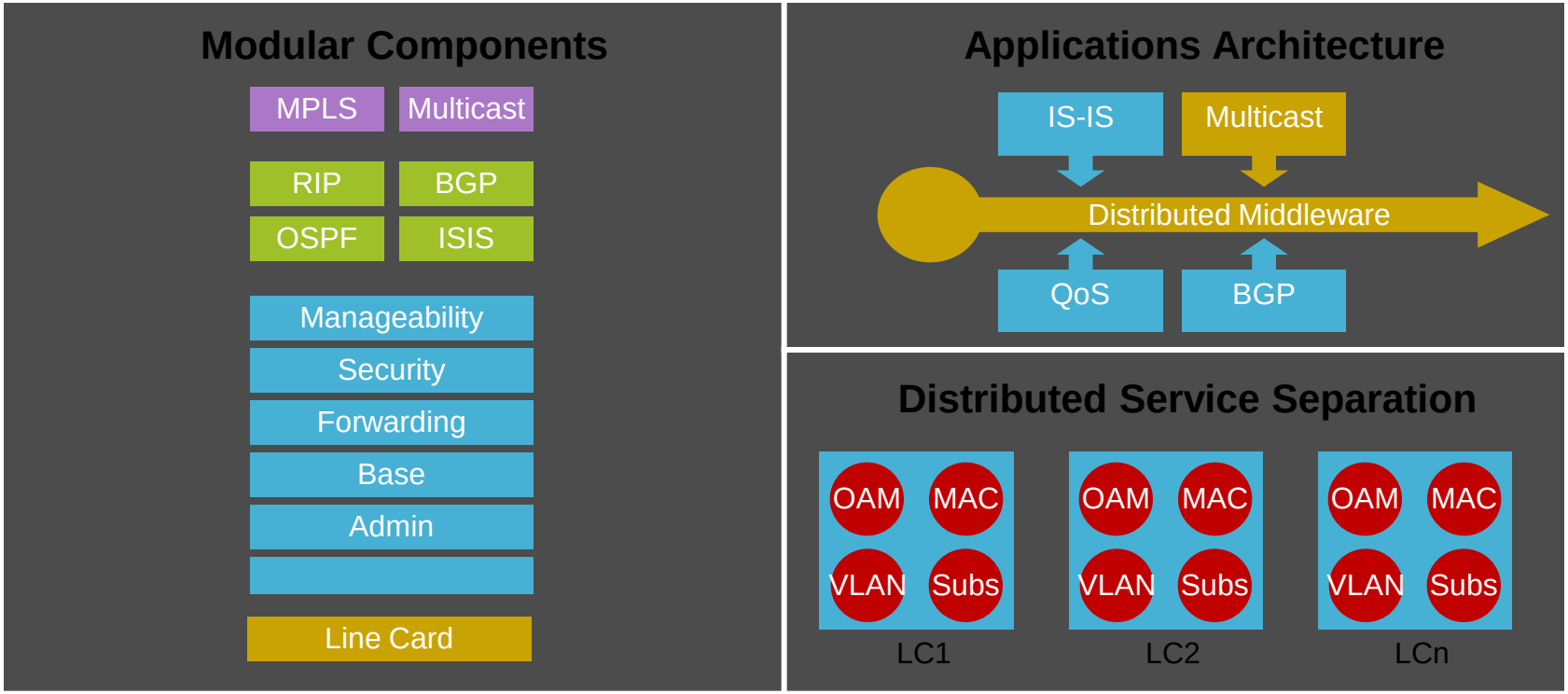
© 2010 Cisco and/or its affiliates. All rights reserved.

Agenda

- **Hardware Architecture**
  - –Chassis
  - –RSP
  - –Line Cards
  - –Switch Fabric
  - –System QoS
  - –Multicast
- **Software Architecture**
  - –IOS-XR and High Availability
  - –L3 Forwarding Architecture
  - –L2 Forwarding Architecture

Carrier Class OS – IOS-XR

A Fully Distributed, Microkernel-Based, True Modular Architecture



Next Generation Architecture

- Microkernel-based design
- Highly modular, highly extensible

Designed for Scale, HA, and Performance

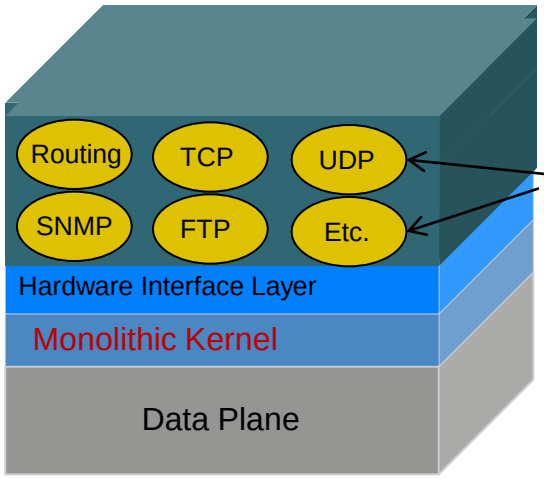
- Scale through distribution
- Unique address tables per linecard
- Process-level, stateful subscriber HA

© 2010 Cisco and/or its affiliates. All rights reserved.

OS – High Level Comparison

Micro-Kernel vs. Monolithic, True Modular vs. Semi-Modular

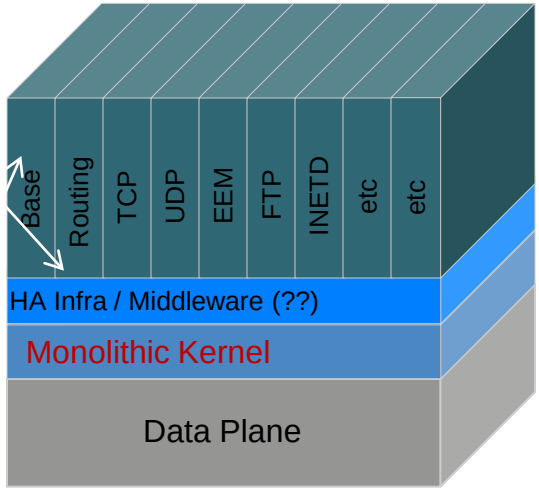
Monolithic OS



Multiple Processes under One Module

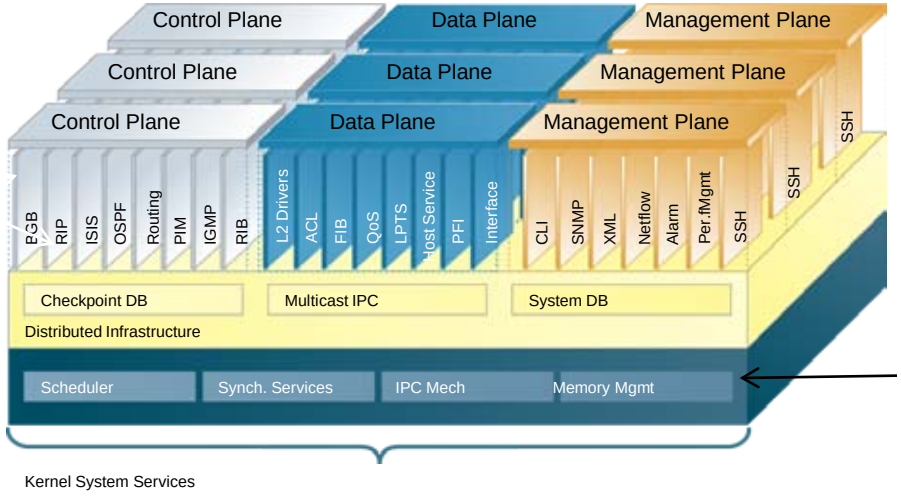
Semi-Modular OS

Restartable Modules



True Modular OS

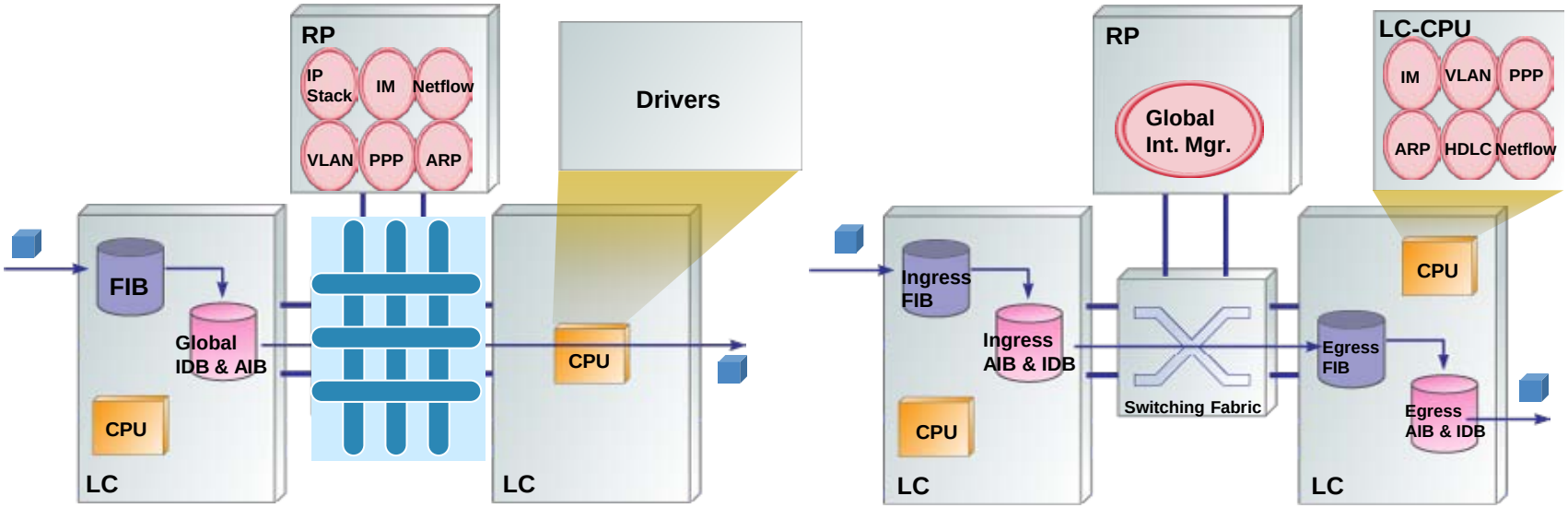
Each Protocol Mapped to one Restartable Module



Micro-Kernel

© 2010 Cisco and/or its affiliates. All rights reserved.
Cisco Public

Fully Distributed Forwarding Infrastructure



Single Stage Forwarding

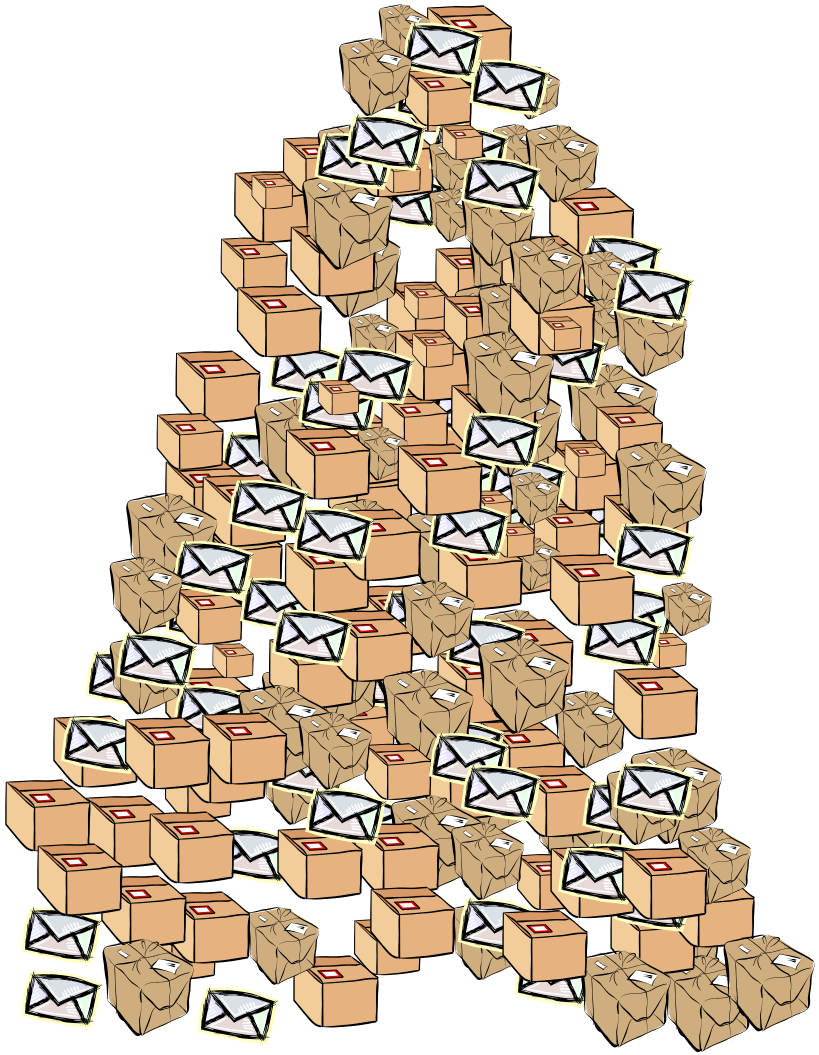
- Single global Adjacency Information Base (AIB) distributed to all line cards
- Single global Interface Management DB distributed to all line cards
- Only Ingress FIB – forces forwarding features to be run in RP

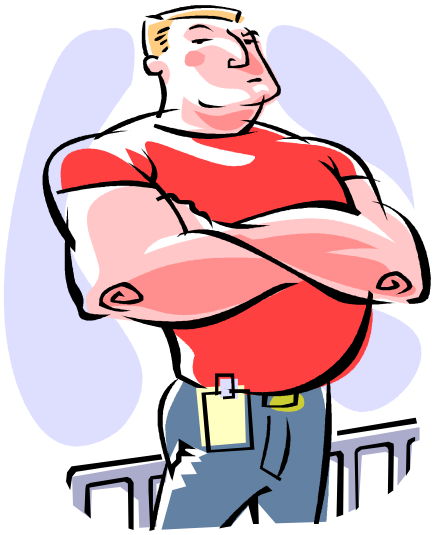
Two Sage Forwarding

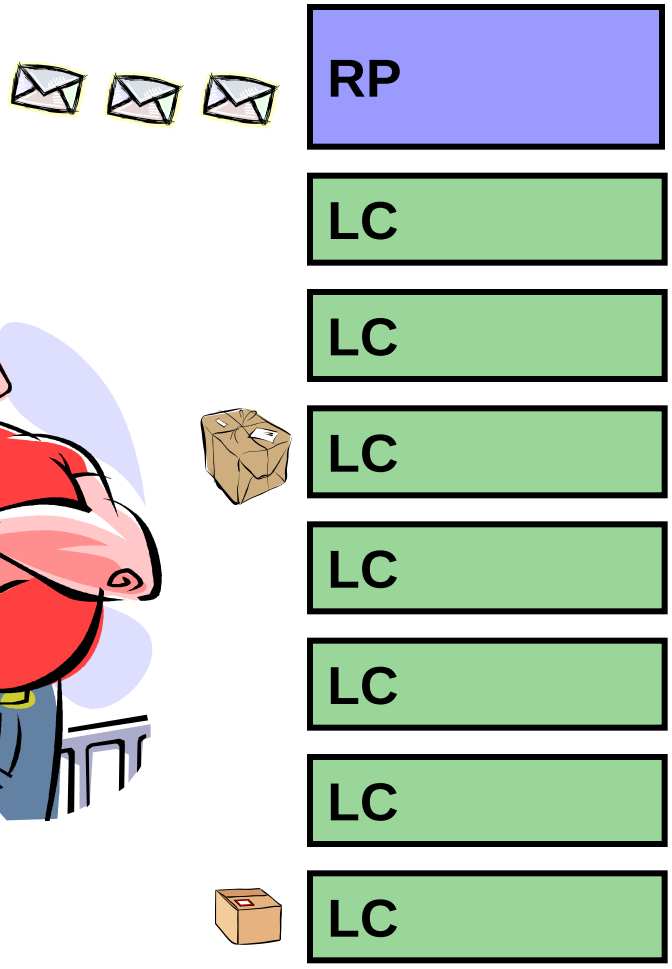
- Each line card has independent AIB only for local interfaces
- Each line card has independent Interface DB for local interfaces
- Both Ingress and Egress FIB – allows forwarding features to be independently run in LCs

© 2010 Cisco and/or its affiliates. All rights reserved.

IOS-XR Control Plane Secure – Protect Against DoS Attack

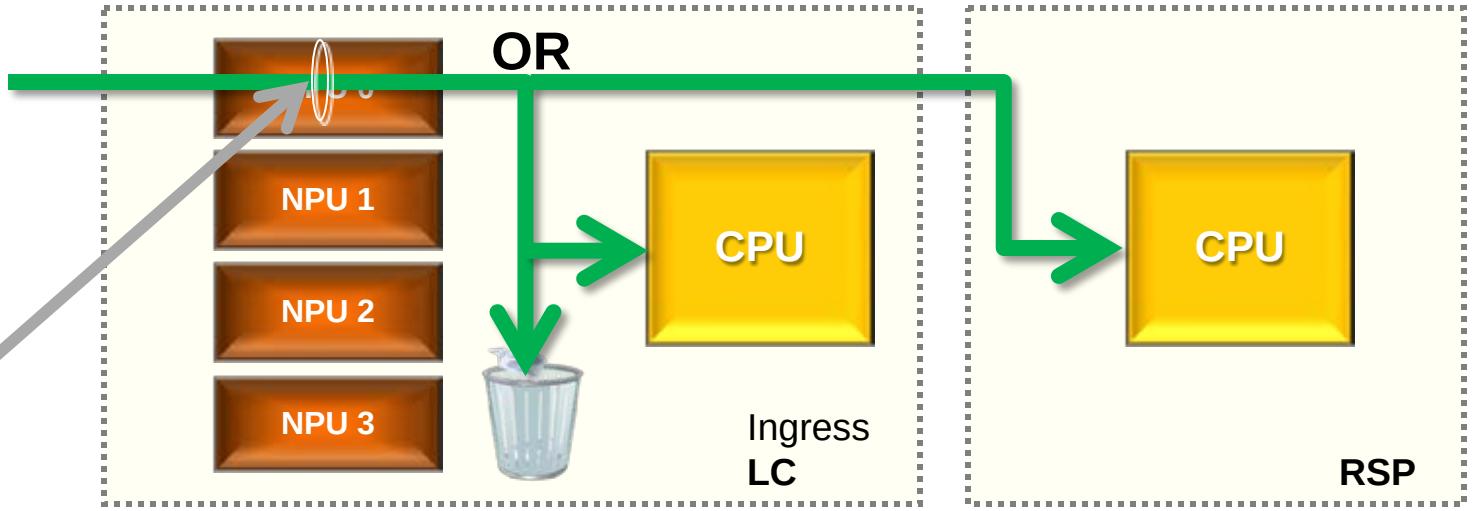






© 2010 Cisco and/or its affiliates. All rights reserved.

IOS-XR LPTS – Secure the Control Plane



- NPU on ingress LC identifies "for us" packets
  - Layer 3 packets use IP FIB
  - Layer 2 frames use various other lookup keys & match criter
- "For us" packets subject to preliminary Internal F determine system destination
  - Pre-IFIB lookup: matches critical packets: routing protocols & L2 control protocols
  - Lookup Result: - Destination: "Local LC CPU", "RSP CPU", or "Drop"
    - Divert priority, queue, and policing rate
- Packets diverted to control plane policed **by the NPU** prior reaching CPU

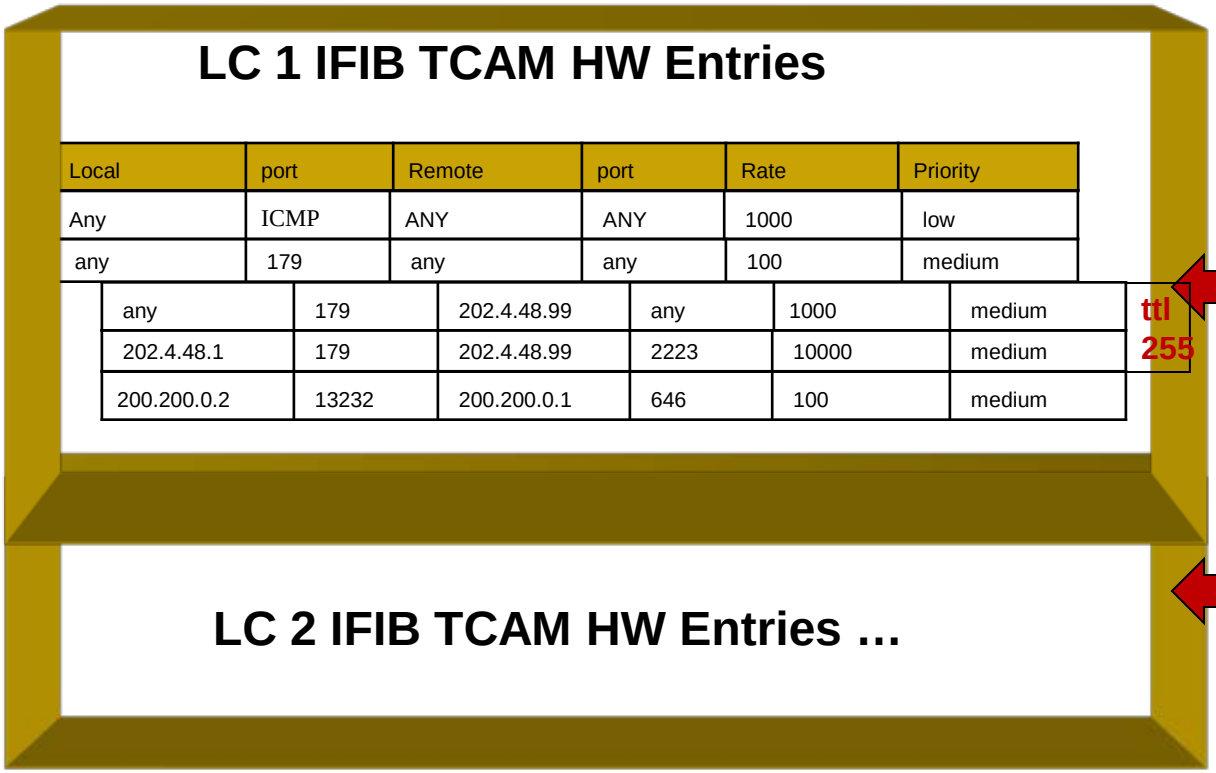
Local Packet Transport Service (LPTS) Policing
- CoPP

© 2010 Cisco and/or its affiliates. All rights reserved. Cisco Public

LPTS – Dynamic Control Plane Protection

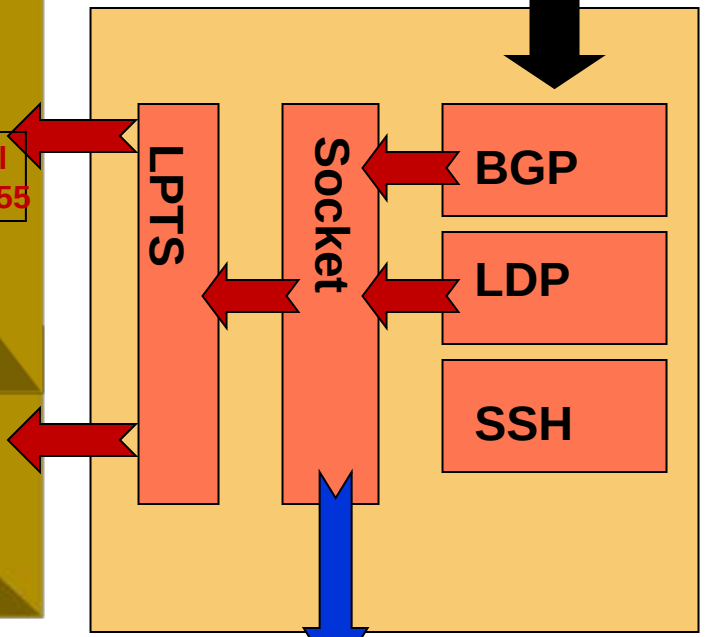
- DCoPP is an automatic, built in firewall for control plane traffic.
- Every Control and Management packet from the line card is rate limited in hardware to provide flood protect at RP

```
router bgp
  neighbor 202.4.48.99
  ttl_security
!
mpls ldp
  ...
!
```



Local	port	Remote	port	Rate	Priority
Any	ICMP	ANY	ANY	1000	low
any	179	any	any	100	medium
any	179	202.4.48.99	any	1000	medium
202.4.48.1	179	202.4.48.99	2223	10000	medium
200.200.0.2	13232	200.200.0.1	646	100	medium

ttl 255



© 2010 Cisco and/or its affiliates. All rights reserved.

LPTS CLI Examples

RP/0/RSP1/CPU0:SJC#show lpts pifib hardware ?

```
  context  Hardware related data structures
  entry    Pre-IFIB hardware entries
  police   Policer values being used
  usage    Hardware table usage
```

RP/0/RSP1/CPU0:SJC(config)#$low telnet default rate ? ← change the default policer rate

<0-4294967295> Packets Per Second

RP/0/RSP1/CPU0:SJC#show lpts pifib hardware police location 0/0/cpu0

```
<snip>
FlowType                Policer Type    Cur. Rate  Def. Rate  Accepted             Dropped
----------------------- ------- ------- ---------- ---------- -------------------- ---------
   -----------
<snip>
OSPF-uc-known           104     Static  1000       1000       0                    0
OSPF-uc-default         105     Static  2000       2000       0                    0
ISIS-known              143     Static  1500       1500       0                    0
ISIS-default            144     Static  2000       2000       0                    0
BGP-known               106     Static  1500       1500       43127                0
<snip>
```


© 2010 Cisco and/or its affiliates. All rights reserved. Cisco Public

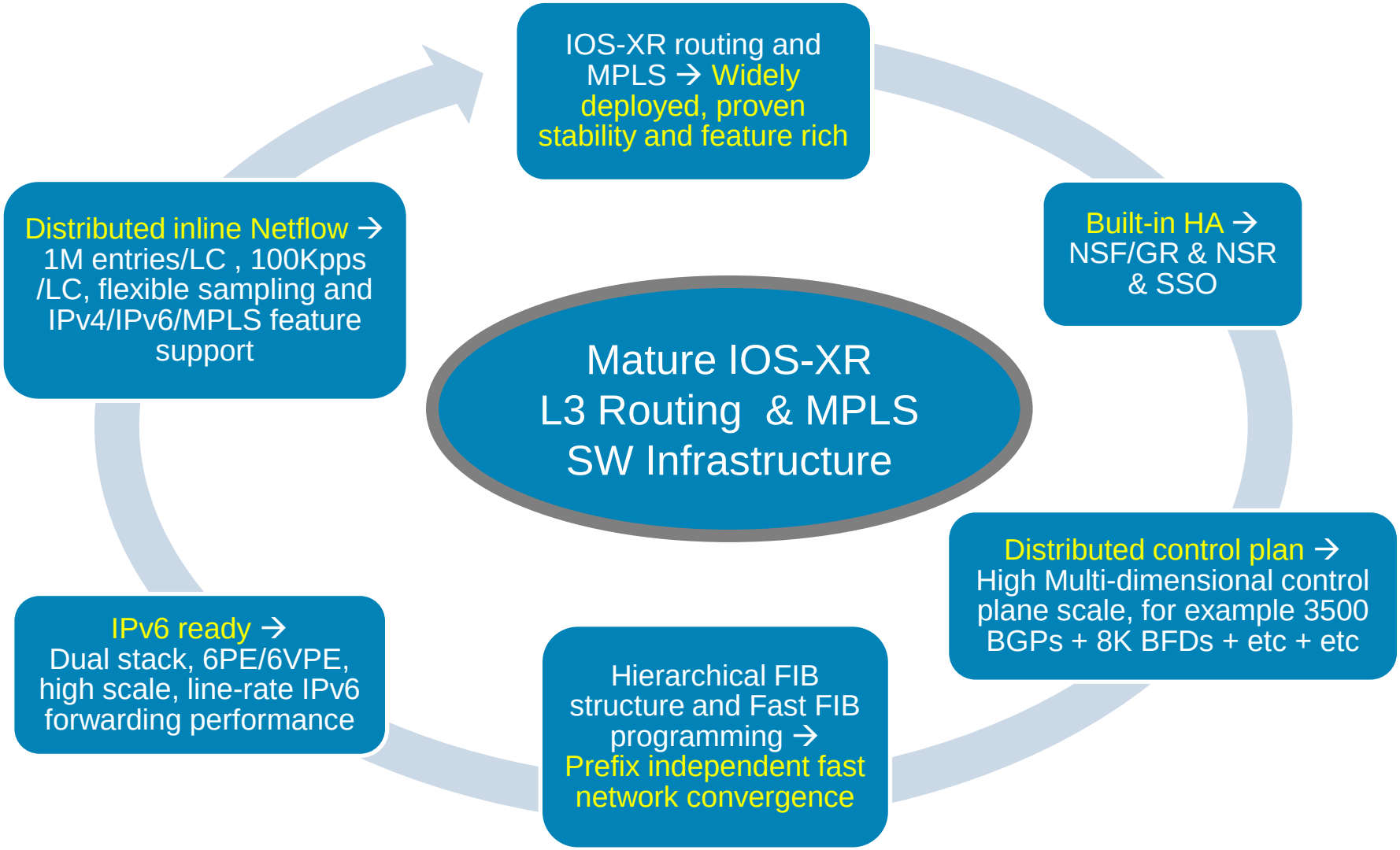
Agenda

- **Hardware Architecture**
  - –Chassis
  - –RSP
  - –Line Cards
  - –Switch Fabric
  - –System QoS
  - –Multicast
- **Software Architecture**
  - –IOS-XR and High Availability
  - –L3 Forwarding Architecture
  - –L2 Forwarding Architecture

© 2010 Cisco and/or its affiliates. All rights reserved.

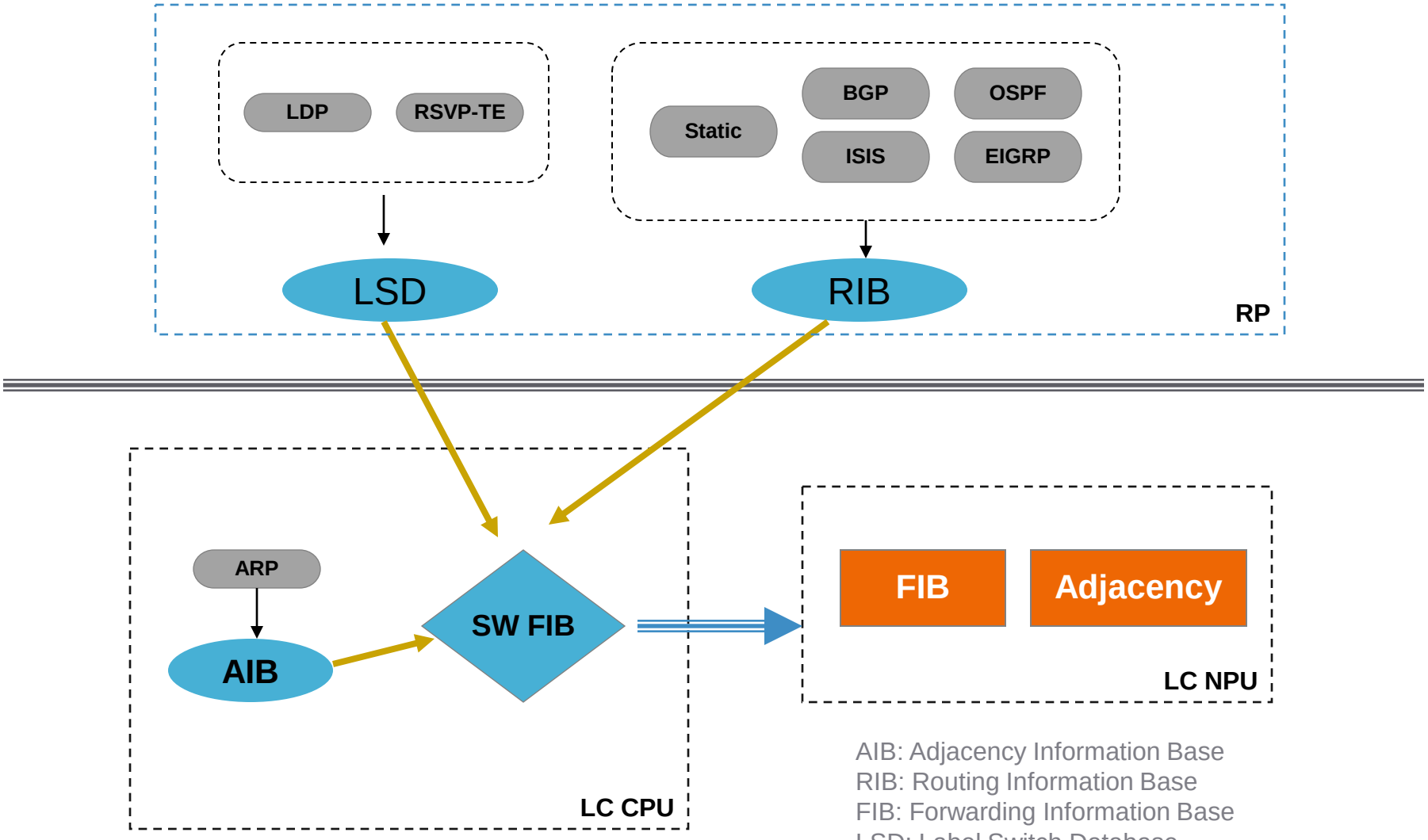
ASR 9000 L3 “At-a-Glance”

Inherit IOS-XR L3/MPLS SW Infrastructure


IOS-XR routing and MPLS → Widely deployed, proven stability and feature rich
Built-in HA → NSF/GR & NSR & SSO
Distributed control plan → High Multi-dimensional control plane scale, for example 3500 BGPs + 8K BFDs + etc + etc
Hierarchical FIB structure and Fast FIB programming → Prefix independent fast network convergence
IPv6 ready → Dual stack, 6PE/6VPE, high scale, line-rate IPv6 forwarding performance
Distributed inline Netflow → 1M entries/LC , 100Kpps /LC, flexible sampling and IPv4/IPv6/MPLS feature support
Mature IOS-XR L3 Routing & MPLS SW Infrastructure

© 2010 Cisco and/or its affiliates. All rights reserved.

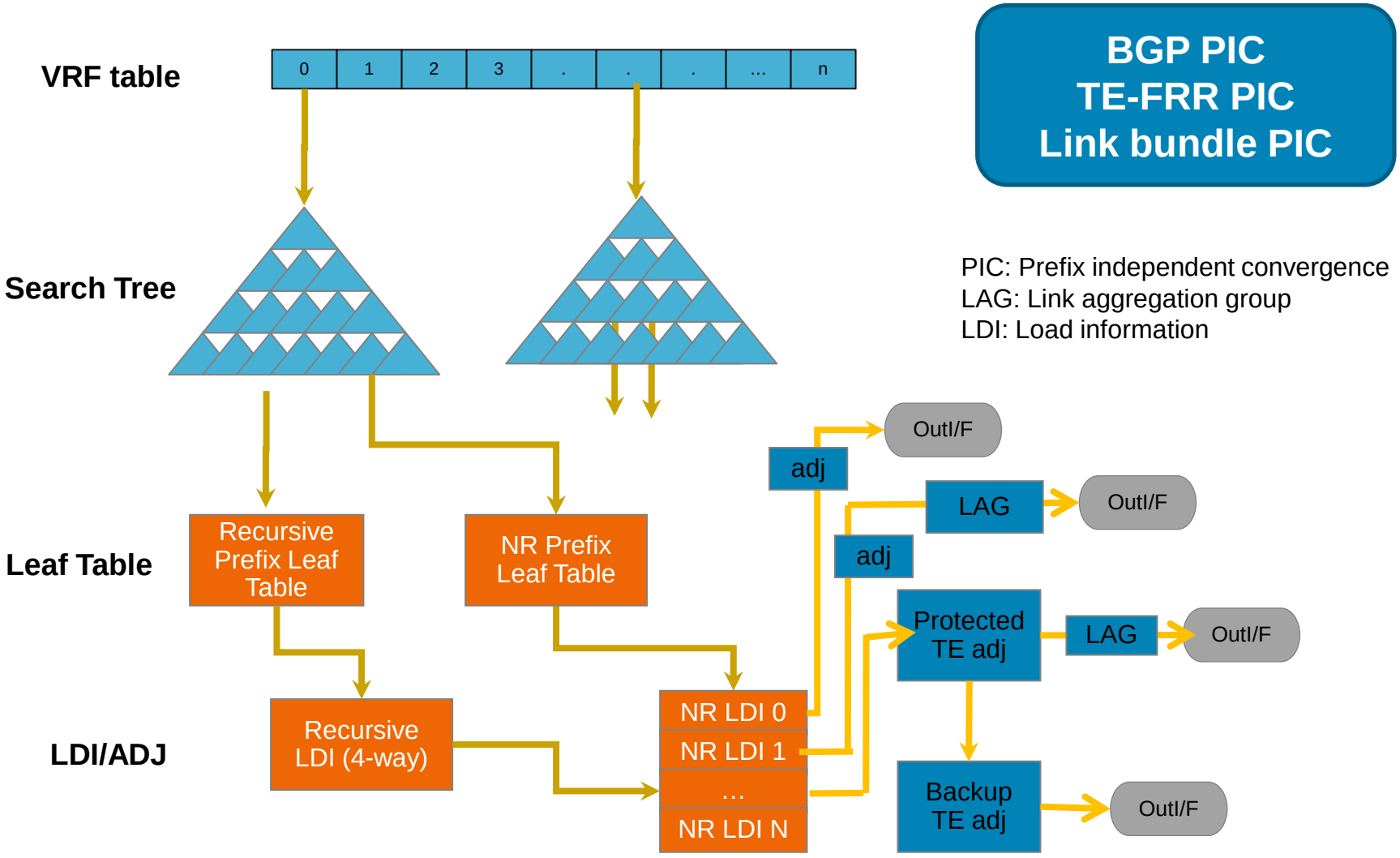
L3 Control Plane Overview



AIB: Adjacency Information Base
RIB: Routing Information Base
FIB: Forwarding Information Base
LSD: Label Switch Database

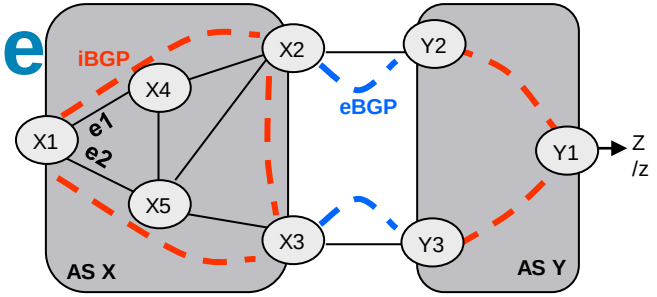
© 2010 Cisco and/or its affiliates. All rights reserved.

L3 Forwarding – Hierarchical Data Structure
FIB and ADJ Tables Overview

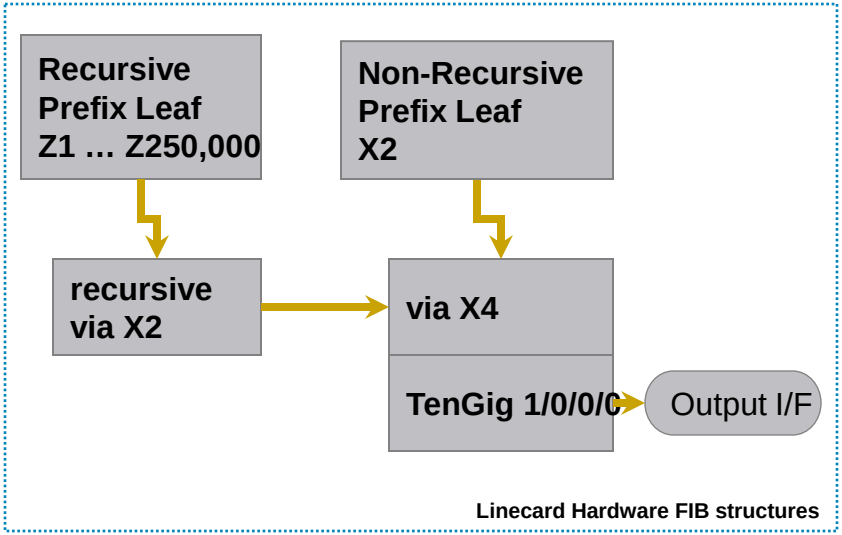


© 2010 Cisco and/or its affiliates. All rights reserved.
Cisco Public

BGP PIC – one simple example

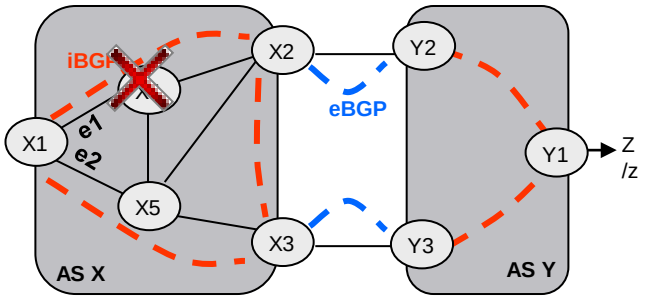


- On X1, it learn 250K routes from BGP. Those 250K routes (Z1…Zn) is in the “recursive” FIB leaf table, since the next hop of the BGP routes is the X2’s loopback address, which is not directly connected
- On X1, it also learn the X2’s loopback address route from IGP. This route is in the “non-recursive” FIB leaf table, since the next hop of this route is the local interface pointing to X4

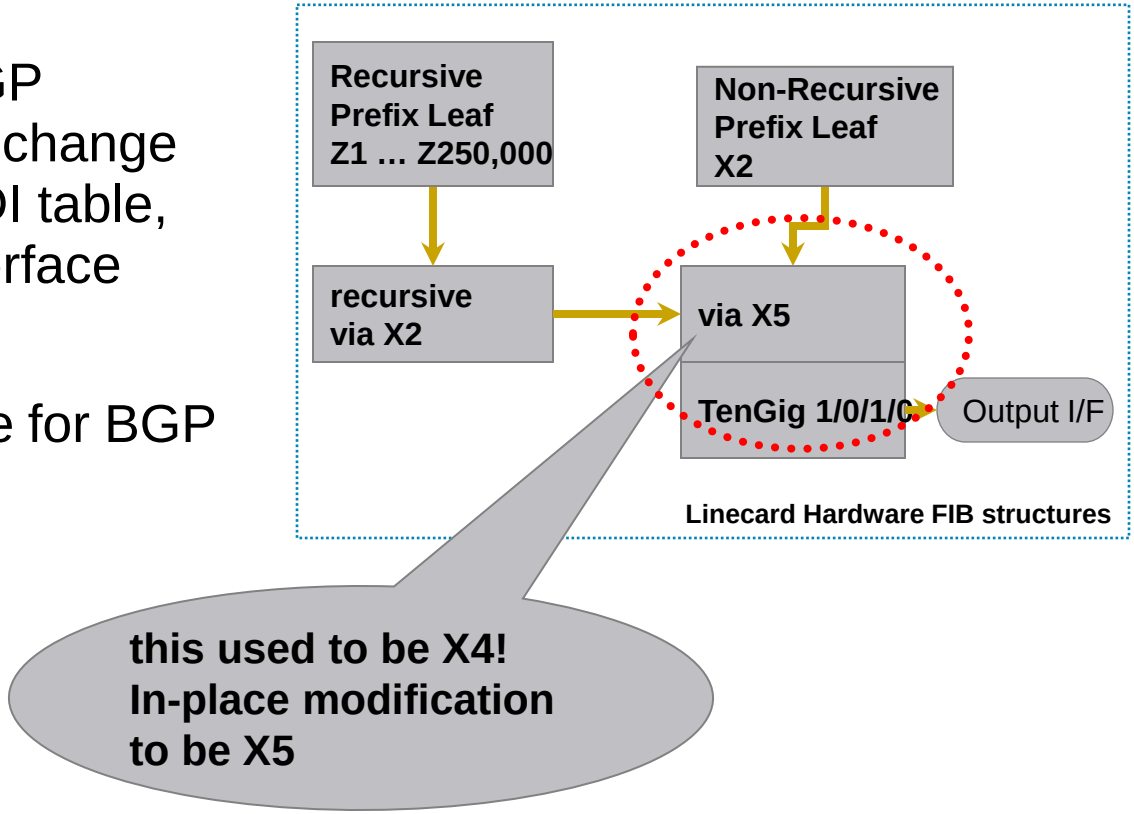


© 2010 Cisco and/or its affiliates. All rights reserved.

BGP PIC – failure example



- If X4 node or X1-X4 link fail …
- IGP converges and now the best IGP path to X2 is via X5
- Instead of writing 250k BGP changes, we only need to change the ONE non-recursive LDI table, which is to replace X4/interface with X5/interface
- O(1) vs. O(n) convergence for BGP



© 2010 Cisco and/or its affiliates. All rights reserved.

ASR 9000 IGP/BGP FC Feature Highlights

- ISIS/OSPF is optimized for FC by default
- IGP Prefix prioritization → guarantee important prefixes get converged at first
- IP FRR → <50msec IP fast reroute without using MPLS TE/FRR
- BGP PIC core → Prefix independent BGP fast convergence for core failure
- BGP PIC Edge → Prefix independent BGP fast convergence for edge failure
- MPLS TE/FRR → <50msec link and node protection, prefix independent
- Link bundle → <50msec link protection for thousands of sub-interfaces per bundle port
- Distributed BFD → 1K session per LC, 8K per system with 150 msec hello interval. Minimal hello interval can be as low as 15msec; BFD support IGP, BGP, static router, PIM, HSRP/VRRP, MPLS TE
- **Multicast only FRR (MoFRR) →**
  - **–50msec IPv4 multicast convergence for IPTV Video service**
  - **–End-to-End 50msec path protection by using data plane failure detection**
  - **–Operational simple without using P2MP TE-FRR**

© 2010 Cisco and/or its affiliates. All rights reserved.

Show CLI Examples – Validate the Forwarding Dath

- show ip route 10.0.0.4
- show cef 10.0.0.4/32 ← RSP's view
- Show cef 10.0.0.4/32 location 0/2/cpu0 ← Local Line card's view
- Show cef 10.0.0.4/32 detail location 0/2/cpu0 ← detailed view on the line card
- show mpls forwarding prefix 10.0.0.4/32
- show arp 10.0.13.3
- show adjacency ten 0/2/0/1 detail location 0/2/cpu0

© 2010 Cisco and/or its affiliates. All rights reserved.

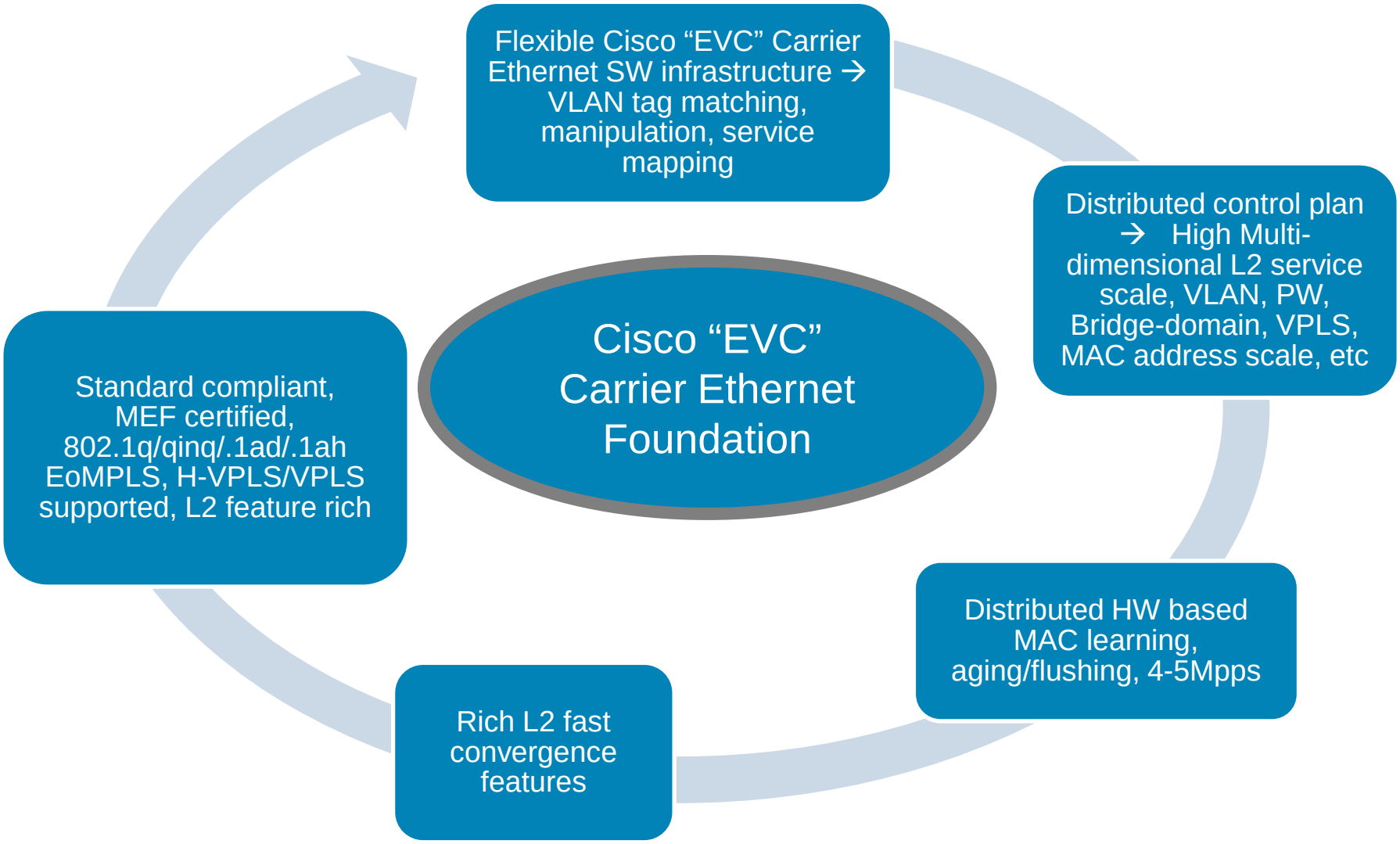
Agenda

- **Hardware Architecture**
  - –Chassis
  - –RSP
  - –Line Cards
  - –Switch Fabric
  - –System QoS
  - –Multicast
- **Software Architecture**
  - –IOS-XR and High Availability
  - –L3 Forwarding Architecture
  - –L2 Forwarding Architecture

© 2010 Cisco and/or its affiliates. All rights reserved.

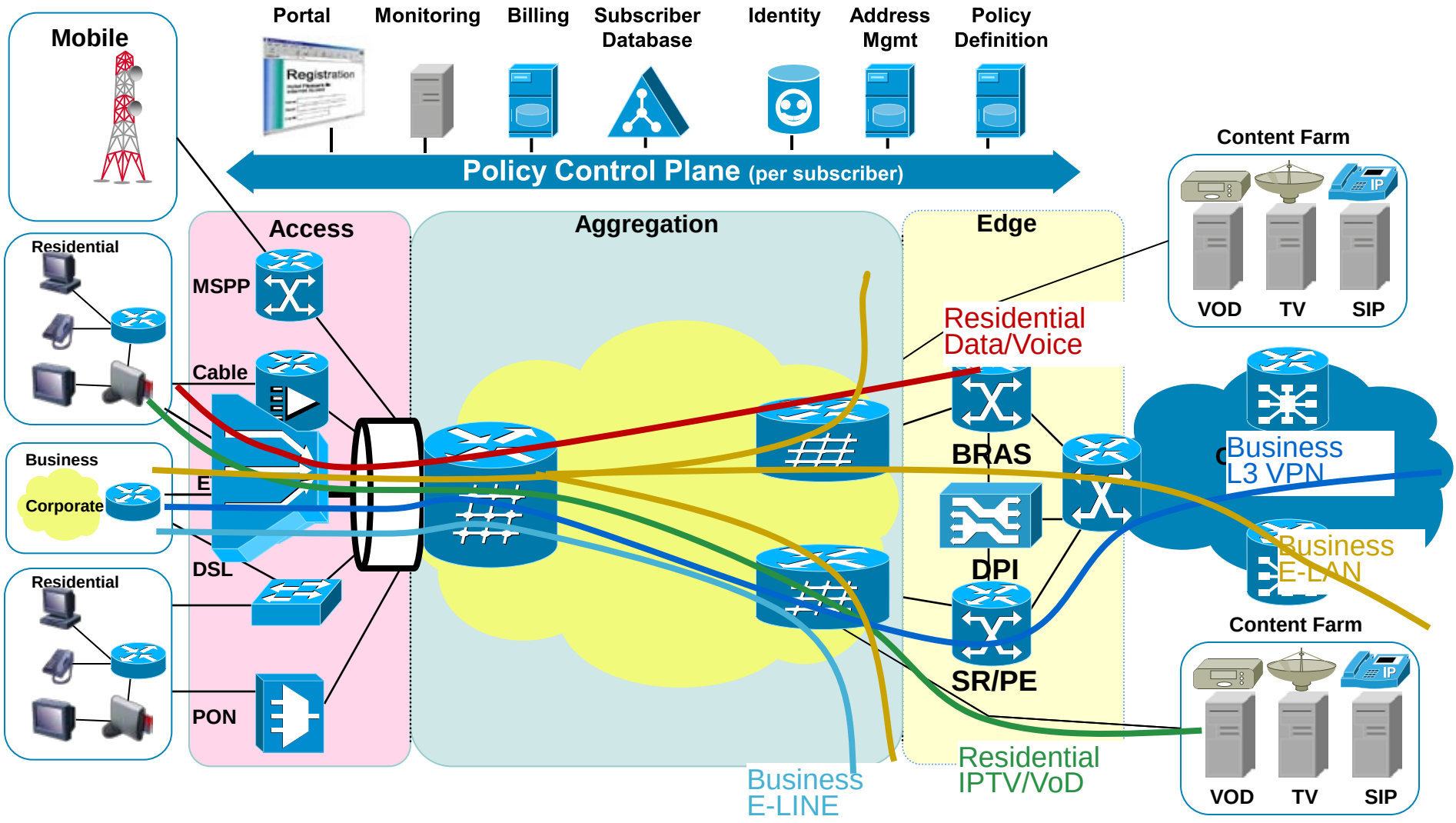
ASR 9000 L2 “At-a-Glance”

Inherit Cisco 7600 Carrier Ethernet L2 SW Foundation


Flexible Cisco “EVC” Carrier Ethernet SW infrastructure → VLAN tag matching, manipulation, service mapping
Distributed control plan → High Multi-dimensional L2 service scale, VLAN, PW, Bridge-domain, VPLS, MAC address scale, etc
Distributed HW based MAC learning, aging/flushing, 4-5Mpps
Rich L2 fast convergence features
Standard compliant, MEF certified, 802.1q/qinq/.1ad/.1ah EoMPLS, H-VPLS/VPLS supported, L2 feature rich
Cisco “EVC” Carrier Ethernet Foundation

© 2010 Cisco and/or its affiliates. All rights reserved.

Converged Edge Network – Flexible Edge Requirements



© 2010 Cisco and/or its affiliates. All rights reserved.

ASR 9000 Flexible Ethernet SW Infrastructure
("EVC" SW Infrastructure)

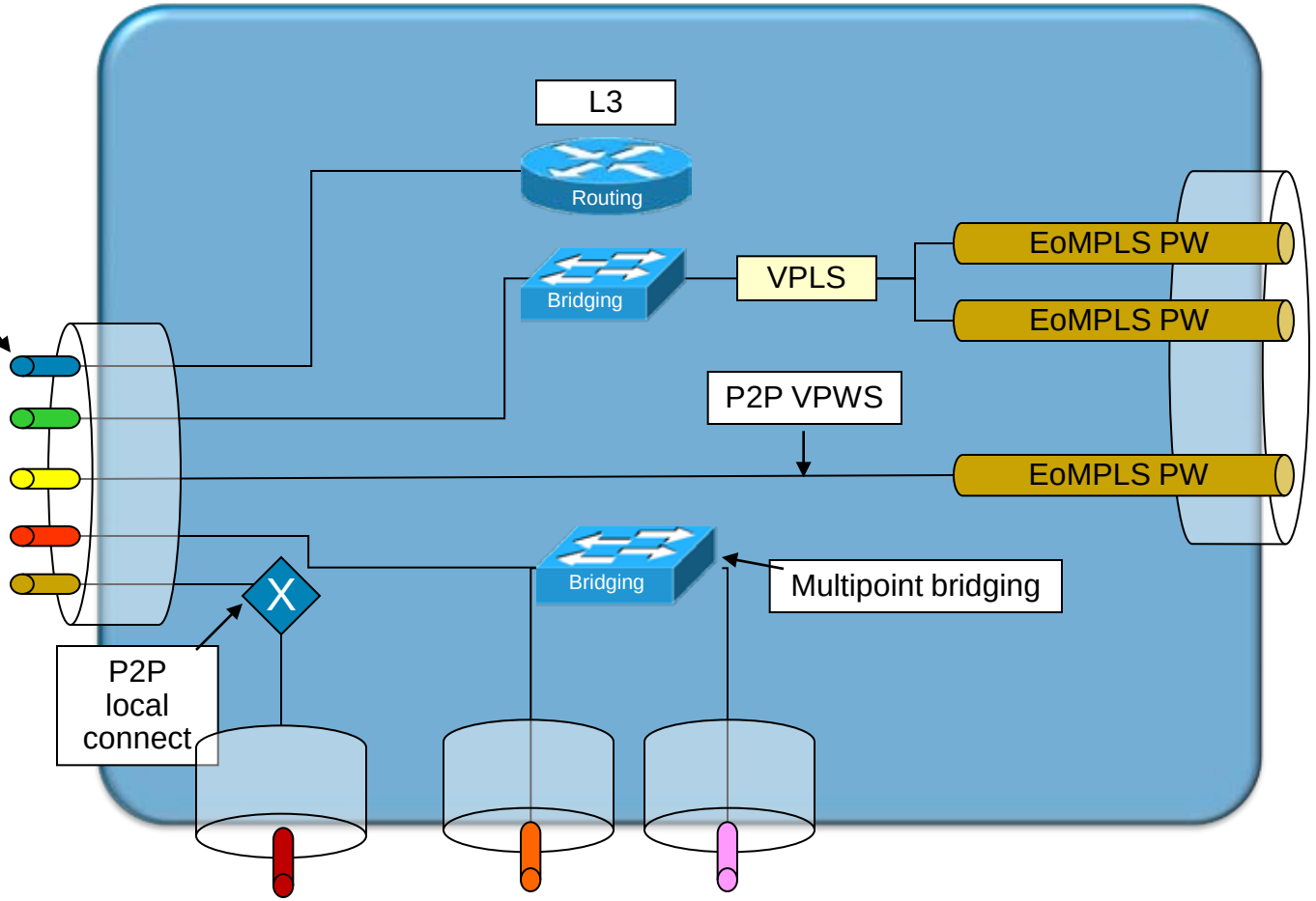
EFP (**E**thernet **F**low **P**oint) or sub-interface

Flexible VLAN tag classification

Flexible VLAN tag rewrite

Flexible Ethertype (.1Q, QinQ, .1ad)

1



2

Flexible service mapping and multiplexing

L2 and L3, P2P and MP services concurrently on the same port

© 2010 Cisco and/or its affiliates. All rights reserved. Cisco Public

Flexible VLAN Tag Classification

```
RP/0/RSP0/CPU0:PE2-asr(config)#int gig 0/0/0/4.100 l2transport
RP/0/RSP0/CPU0:PE2-asr(config-subif)#encapsulation ?
  default   Packets unmatched by other service instances
  dot1ad    IEEE 802.1ad VLAN-tagged packets
  dot1q     IEEE 802.1Q VLAN-tagged packets
  untagged  Packets with no explicit VLAN tag

RP/0/RSP0/CPU0:PE2-asr(config-subif)#encapsulation dot1q 100-200,205 ?
  comma  comma
  exact  Do not allow further inner tags

RP/0/RSP0/CPU0:PE2-asr(config-subif)#encapsulation dot1q 100 second-dot1q 10-20,35 ?
  comma  comma
  exact  Do not allow further inner tags

RP/0/RSP0/CPU0:PE2-asr(config-subif)#encapsulation dot1ad 20 dot1q 10-20 ?
  comma  comma
  exact  Do not allow further inner tags
```

- Double tag
- Single tag
- Multiple tag
- Range of tag
- .1q and/or .1ad
- Loose or exact match
- Untagged
- Default tag

© 2010 Cisco and/or its affiliates. All rights reserved.
Cisco Public

Flexible VLAN Tag Rewrite

```
RP/0/RSP0/CPU0:PE2-asr(config)#int gig 0/0/0/4.100 l2transport

RP/0/RSP0/CPU0:PE2-asr(config-subif)#rewrite ingress tag ?
  pop        Remove one or more tags
  push       Push one or more tags
  translate  Replace tags with other tags

RP/0/RSP0/CPU0:PE2-asr(config-subif)#rewrite ingress tag pop ?
  1  Remove outer tag only
  2  Remove two outermost tags

RP/0/RSP0/CPU0:PE2-asr(config-subif)#rewrite ingress tag push ?
  dot1ad  Push a Dot1ad tag
  dot1q   Push a Dot1Q tag
RP/0/RSP0/CPU0:PE2-asr(config-subif)#rewrite ingress tag push dot1q 100 ?
  second-dot1q  Push another Dot1Q tag
  symmetric     All rewrites must be symmetric

RP/0/RSP0/CPU0:PE2-asr(config-subif)#rewrite ingress tag translate ?
  1-to-1  Replace the outermost tag with another tag
  1-to-2  Replace the outermost tag with two tags
  2-to-1  Replace the outermost two tags with one tag
  2-to-2  Replace the outermost two tags with two other tags
```

Pop tag 1 or 2

Push tag 1 or 2

Tag translation

- 1-1
- 1-2
- 2-1
- 2-2

© 2010 Cisco and/or its affiliates. All rights reserved.

Flexible Service – L2VPN P2P

EFP configuration example

```
Interface gig 0/0/0/1.101 l2transport
encapsulation dot1q 101 second 10
rewrite ingress pop 2 Symmetric

Interface gig 0/0/0/2.101 l2transport
encapsulation dot1q 101
rewrite ingress pop 1 Symmetric

Interface gig 0/0/0/3.101 l2transport
encapsulation dot1q 102
rewrite ingress push dot1q 100 Symmetric
```

L2VPN P2P service configuration example

```
l2vpn
 xconnect group cisco
   p2p service1 ← local connect
     interface gig 0/0/0/1.101
     interface gig 0/0/0/2.101
   p2p service2 ← VPWS
     interface gig 0/0/0/3.101
     neighbor 1.1.1.1 pw-id 22
  p2p service3 ← PW stitching
     neighbor 2.2.2.2 pw-id 100
    neighbor 3.3.3.3 pw-id 101
```


© 2010 Cisco and/or its affiliates. All rights reserved.

Flexible Service – L2VPN Multi-Point

EFP configuration example

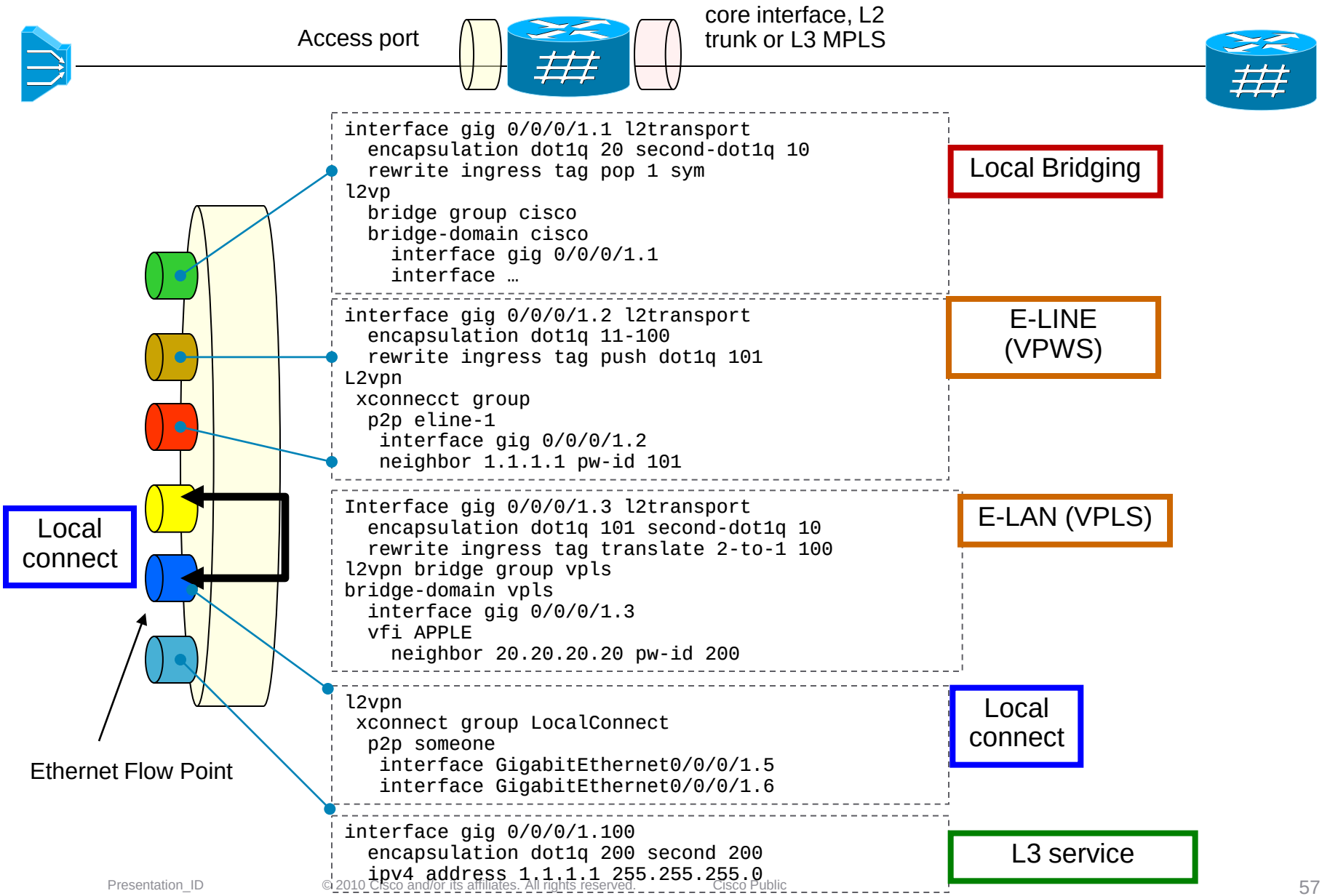
```
Interface gig 0/0/0/1.101 l2transport
encapsulation dot1q 101
rewrite ingress pop 1 Symmetric

Interface gig 0/0/0/2.101 l2transport
encapsulation dot1q 101
rewrite ingress pop 1 Symmetric

Interface gig 0/0/0/3.101 l2transport
encapsulation dot1q 102
rewrite ingress push dot1q 100 Symmetric
```

L2VPN MP service configuration example

```
l2vpn
 bridge group cisco
  bridge-domain domain1 ← local bridging
   Interface gig 0/0/0/1.101
     split-horizon group ← no bridging among same SHG
   Interface gig 0/0/0/2.101
     split-horizon group

  bridge-domain domain2 ← vpls
   Interface gig 0/0/0/1.101
   Interface gig 0/0/0/2.101
   vfi cisco
    neighbor 192.0.0.1 pw-id 100
    neighbor 192.0.0.2 pw-id 100

  bridge-domain domain3 ← h-vpls
   Interface gig 0/0/0/1.101
   neighbor 192.0.0.3 pw-id 100 ← spoke PW
   vfi cisco ← for core PWs
    neighbor 192.0.0.1 pw-id 100 ← core PW
    neighbor 192.0.0.2 pw-id 100
```


© 2010 Cisco and/or its affiliates. All rights reserved.

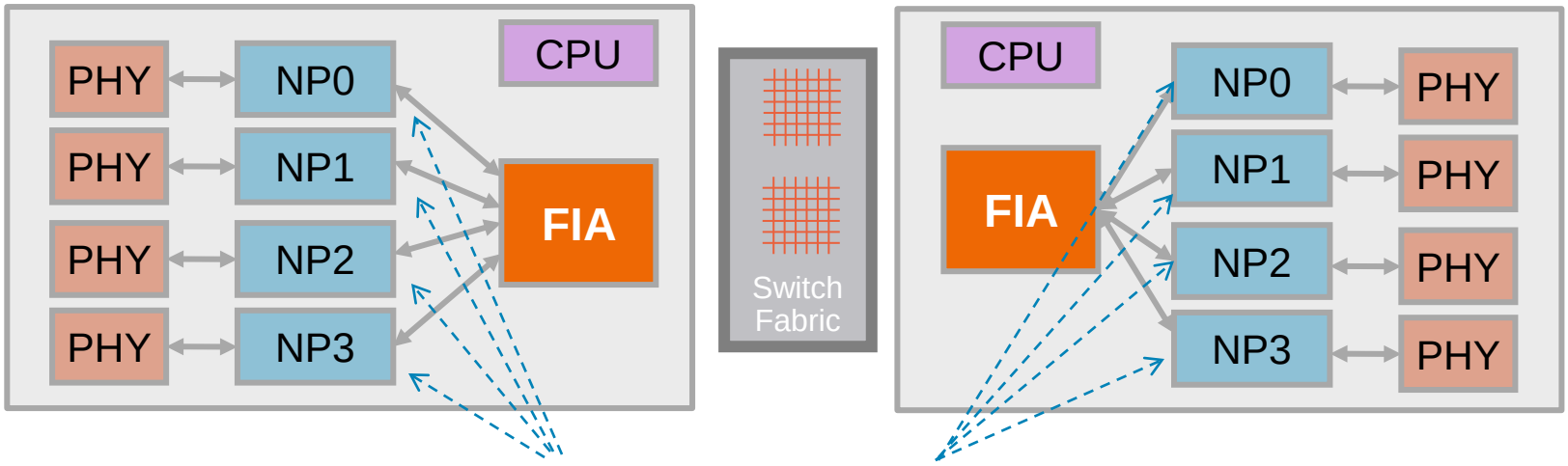
Multiple Services on the same port example



© 2010 Cisco and/or its affiliates. All rights reserved.
Cisco Public

ASR 9000 MAC Learning Overview

- **Distributed HW based MAC Learning**
  - –MAC learning /aging/flushing is fully distributed to the NP, independently
  - –Each NP can learn around 4-5 Millions MAC addresses per second in hardware
- **Data-plane MAC table synchronization**
  - –MAC address are synchronized across all NPs in the entire system, regardless of the bridge-domain or bridge port configuration.
  - –MAC address synchronization is achieved by special MAC notification messages which is implemented in the data-plane microcode



MAC learning/aging/flushing is done by hardware and fully distributed on each NP independently

L2 Fast Convergence Feature Highlights

- PW Redundancy → EoMPLS service level redundancy
- H-VPLS redundancy with MAC withdrawal → (H-)VPLS service redundancy
- L2VPN MPLS TE/FRR → <50msec link and node protection, PWs independent
- Link bundle → <10msec link protection for thousands of L2 sub-interfaces per bundle port
- Distributed CFM → 32K CFM session per LC, 100msec CFM CC timer
- MST access gateway → Access network redundancy, access network run MST
- REP access gateway → Access network redundancy, access network run REP
- Multi-chassis LAG (4.0 release, 2HCY10)
- MST → native L2 network
- G8032, G8031 (target for future release)

© 2010 Cisco and/or its affiliates. All rights reserved.

CLI Comparison (1) – L2VPN P2P service

	ASR 9000	7600
Local Connect	**EFP configuration under interface** **Including VLAN tag encapsulation, tag rewrite, Qo/ACL features, etc** Interface gig 0/0/0/1.101 l2transport encapsulation dot1q 101 second 10 rewrite ingress tag pop 2 Symmetric Interface gig 0/0/0/2.101 l2transport encapsulation dot1q 101 rewrite ingress tag pop 1 Symmetric	interface GigabitEthernet4/1/0 service instance 101 ethernet encapsulation dot1q 101 second 10 rewrite ingress tag pop 2 Symmetric interface GigabitEthernet4/1/1 service instance 100 ethernet encapsulation dot1q 100 rewrite ingress tag pop 1 Symmetric connect eline-101 GigabitEthernet4/1/0 101 GigabitEthernet4/1/1 100
EoMPLS	**Service configuration under "l2vpn"** l2vpn xconnect group cisco p2p service1 ← local connect interface gig 0/0/0/1.101 interface gig 0/0/0/2.101 p2p service2 ← EoMPLS interface gig 0/0/0/3.101 neighbor 1.1.1.1 pw-id 22 p2p service3 ← PW stitching neighbor 2.2.2.2 pw-id 100 neighbor 3.3.3.3 pw-id 101	interface GigabitEthernet4/1/1 service instance 11 ethernet encapsulation dot1q 101 second-dot1q 60-70 xconnect 10.0.0.3 101 encapsulation mpls
PW stitching		l2 vfi tac-training point-to-point neighbor 10.0.2.3 3001 encapsulation mpls neighbor 10.0.2.2 3000 encapsulation mpls [note] require BGP configuration if it's for inter-AS

© 2010 Cisco and/or its affiliates. All rights reserved.

CLI Comparison (2) – L2VPN MP bridging

	ASR 9000	7600
Local Bridging	**EFP configuration under interface** **Including VLAN tag encapsulation, tag rewrite, Qo/ACL features, etc** Same as L2VPN P2P services **Service configuration under "l2vpn"**	interface GigabitEthernet4/1/0 service instance 101 ethernet encapsulation dot1q 101-1000 **bridge-domain 100 split-horizon** interface GigabitEthernet4/1/1 service instance 101 ethernet encapsulation dot1q 101-1000 **bridge-domain 100 split-horizon**
(H-)VPLS & SVI based EoMPLS	l2vpn bridge group cisco bridge-domain domain1 ← local bridging Interface gig 0/0/0/1.101 split-horizon group Interface gig 0/0/0/2.101 split-horizon group bridge-domain domain2 ← SVI EoMPLS Interface gig 0/0/0/1.101 Interface gig 0/0/0/2.101 neighbor 192.0.0.1 pw-id 100 bridge-domain domain3 ← vpls & h-vpls Interface gig 0/0/0/1.101 neighbor 192.0.0.3 pw-id 100 ← spoke PW vfi cisco ← core PWs under VFI config neighbor 192.0.0.1 pw-id 100 neighbor 192.0.0.2 pw-id 100	interface GigabitEthernet4/1/0 service instance 2 ethernet encapsulation dot1q 20 bridge-domain 20 split-horizon interface GigabitEthernet4/1/1 service instance 2 ethernet encapsulation dot1q 20 bridge-domain 20 split-horizon l2 vfi vpls-20 manual ← VPLS vpn id 120 neighbor 10.0.0.1 encapsulation mpls ← core PW neighbor 10.0.0.2 encapsulation mpls no-split-horizon ← spoke PW neighbor 10.0.0.4 encapsulation mpls interface Vlan20 ← (H-)VPLS xconnect vfi vpls-20 interface Vlan20 ← SVI EoMPLS xconnect 10.0.0.1 101 encap mpls

CLI Comparison (3) – “L2 Switchport” Configuration

ASR 9000 CLI example

```
Interface gig 0/0/0/1.50 l2transport
encapsulation dot1q 50
rewrite ingress tag pop 1 Symmetric

Interface gig 0/0/0/1.51 l2transport
encapsulation dot1q 51
rewrite ingress tag pop 1 Symmetric
…
Interface gig 0/0/0/1.1000 l2transport
encapsulation dot1q 1000
rewrite ingress tag pop 1 Symmetric


l2vpn
 bridge group cisco

bridge-domain domain50
   Interface gig 0/0/0/1.50
   Interface gig 0/0/0/2.50
<snip>
bridge-domain domain1000
   Interface gig 0/0/0/1.1000
   Interface gig 0/0/0/2.1000
```

Switchport trunk configuration

```
interface gig 1/1
  switchport
  switchport mode trunk
  switchport trunk allow vlan 50-1000

interface gig 1/2
  switchport
  switchport mode trunk
  switchport trunk allow vlan 50-1000
```

- L2 sub-interface model is more attractive due to the per service features. However, traditional switchport trunk configuration is easier

© 2010 Cisco and/or its affiliates. All rights reserved.

Complete Your Online Session Evaluation

- Give us your feedback and you could win fabulous prizes. Winners announced daily.
- Receive 20 Cisco Preferred Access points for each session evaluation you complete.
- Complete your session evaluation online now (open a browser through our wireless network to access our portal) or visit one of the Internet stations throughout the Convention Center.

Don’t forget to activate your Cisco Live and Networkers Virtual account for access to all session materials, communities, and on-demand and live activities throughout the year. Activate your account at any internet station or visit www.ciscolivevirtual.com.

© 2010 Cisco and/or its affiliates. All rights reserved.

© 2010 Cisco and/or its affiliates. All rights reserved.